CONFERENCE SERIES

Philipp Zech, Justus Piater (Eds.)

Proceedings of the Austrian Robotics Workshop 2018

innsbruck university press

CONFERENCE SERIES

innsbruck university press

universität
innsbruck

Philipp Zech, Justus Piater (Eds.)

Proceedings of the Austrian Robotics Workshop 2018

Philipp Zech
Justus Piater
Department of Computer Science, Universität Innsbruck

Universität Innsbruck
1st edition

www.uibk.ac.at/iup
ISBN 978-3-903187-22-1
DOI 10.15203/3187-22-1

Program Committee

- Mathias Brandstötter – Joanneum Research
- Bernhard Dieber – Joanneum Research
- Martin Humenberger – NAver Labs Europe
- Brigitte Krenn – Austrian Research Institute for Artificial Intelligence
- Gernot Kronreif – Technikum Wien
- Wilfried Kubinger – Technikum Wien
- Andreas Müller – Johannes Keppler University
- Kurt Niel – FH Wels
- Justus Piater – University of Innsbruck
- Andreas Pichler – Profactor GmbH
- Athansios Polydoros – University of Innsbruck
- Erwan Renaudo – University of Innsbruck
- Safoura Rezapour-Lakani – University of Innsbruck
- Bernhard Rinner – Klagenfurt University
- Lukas Silberbauer – Taurob
- Gerald Steinbauer – Techinical University Graz
- Markus Vincze – Technical University Vienna
- Christian Wögerer – Profactor GmbH
- Philipp Zech – University of Innsbruck

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
 ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Content

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
 ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Preface

Philipp Zech and Justus Piater
Department of Computer Science, University of Innsbruck

The 6th Austrian Robotics Workshop of the Austrian Association for Measurement, Automation and Robotics took place May 17–18, 2018, in the Kaiser- Leopold Hall of the University of Innsbruck, Austria, and was attended by 29 participants.

The program was composed of 3 keynote talks by high-profile researchers from outside of Austria, 9 contributed talks, and 3 posters. The contributed talks were selected from 11 submitted articles by peer review. Each article was reviewed by two members of the program committee.

A Best Research Paper award sponsored by the IEEE RAS Austria Section was presented to Florian Pucher for the paper

> Florian Pucher, Hubert Gattringer and Andreas Müller, *Analysis of Feature Tracking Methods for Vision-Based Vibration Damping of Flexible Link Robots*

A Best Student Paper award sponsored by the ABB-Group was presented to Florian Dannereder for the paper

> Florian Dannereder, Paul Herwig Pachschwöll, Mohamed Aburaia, Erich Markl, Maximilian Lackner, and Corinna Engelhardt-Nowitzki, *Development of a 3D-Printed Bionic Hand with Muscle- and Force Control*

A Best Student Poster award sponsored by the GMAR-Robotics was pre- sented to Florian Dannereder for the paper

> Matthias Hirschmanner, Stephanie Gross, Brigitte Krenn, Friedrich Neubarth, Martin Trapp, Michael Zillich, Markus Vincze, *Extension of the Action Verb Corpus for Supervised Learning*

The best papers and poster were selected by the conference chairs and the representatives of the GMAR-Robotics who were present, based on the reviews and the presentations:

- Mathias Brandstötter, Joanneum Research
- Wilfried Kubinger, FH Technikum Wien
- Justus Piater, Universität Innsbruck
- Markus Vincze, TU Wien
- Philipp Zech, Universität Innsbruck

The ARW 2018 Chairs,
Philipp Zech and Justus Piater

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Keynote Speakers

Tamim Asfour (KIT): Engineering Humanoids for the Real World

Abstract — The talk addresses recent progress towards building integrated. 24/7 humanoid robots able to perform complex grasping and manipulation tasks and to learn from human observation and sensorimotor experience. I will present recent regarding the development and applications of humanoid robots in household as well as industrial environments as collaborative robots which provide help for humans. Further, I will address the important questions of motion generation in high dimensional spaces and how learning from human observation and natural language methods can be combined to build a motion alphabet and robot internet of skills as the basis for intuitive and flexible robot programming. I will conclude with a discussion of current development in the area of AI and the challenges of a Robotics AI.

Biography — Tamim Asfour is full Professor of Humanoid Robotics at the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT) where he is head of the High Performance Humanoid Technologies Lab (H2T). His research interest is high performance 24/7 humanoid robotics. Specifically, his research focuses on engineering humanoid robot systems, which are able to perform grasping and dexterous manipulation tasks, learn from human observation and sensorimotor experience as well as on the mechano-informatics of humanoids as the synergetic integration of mechatronics, informatics and artificial intelligence methods into integrated humanoid robot systems. He is developer of the ARMAR humanoid robot family and is leader of the humanoid research group at KIT since 2003. In his research, he is reaching out and connecting to neighboring areas in large-scale national and European interdisciplinary projects in the area of robotics in combination with machine learning and computer vision.

He is the Founding Editor-in-Chief of the IEEE-RAS Humanoids Con- ference Editorial Board, co-chair of the IEEE-RAS Technical Committee on Humanoid Robots (2010-2014), Editor of the Robotics and Automation Letters, Associate Editor of Transactions on Robotics (2010-2014). He is president of the Executive Board of the German Robotics Society (DGR), member of the Board of Directors of euRobotics (2013-2015) and scientific spokesperson of the KIT Center ”Information · Systems · Technologies (KCIST)”.

Stephane Doncieux (ISIR): Open-ended Learning and Development in Robotics

Abstract — Autonomous robots still have a hard time in non-controlled conditions. One of the main reasons is their lack of adaptivity: the robot programmer needs to analyse the task and the environment the robot will have to deal with to design its morphology and define its behavior that will remain the same for the whole robot life. If a situation occurs that has not been foreseen and if the designed behavior cannot deal with it, the robot will fail. Building robots able to deal with such unforeseen situations requires for the robot to go beyond the knowledge it has been endowed with. These robots need to have open-ended learning abilities, i.e. the ability to turn the problem they are face with in a such a way that they can solve it through learning. This implies to to be able to bootstrap skill acquisition with no task specific knowledge and to build adapted representations of state and action spaces so that learning can occur. We present the work done in this direction in the frame of the DREAM European project (http://robotsthatdream.eu/).

Biography — Stephane Doncieux is Professor in Computer Science at the Sorbonne University, in the ISIR lab, in Paris, France. He is responsible of the AMAC research team (Architectures and Models of Adaptation and Cognition). His goal is to design algorithms that allow robots to deal with open environments. His work focuses on evolutionary learning approaches in robotics (Evolutionary Robotics) and Developmental Robotics. He currently focuses his research on how to bootstrap a cognitive robot by allowing it to discover its environment and the objects it contains through its interactions. This question, centered on the ability to acquire experience and restructure the representations the robots relies on, is the central topic of the DREAM European project (FET H2020), that he coordinates (http://robotsthatdream.eu/).

Jan Peters (TU Darmstadt): Robot Skill Learning

Abstract — Autonomous robots that can assist humans in situations of daily life have been a long standing vision of robotics, artificial intelligence, and cognitive sciences. A first step towards this goal is to create robots that can learn tasks triggered by environmental context or higher level in- struction. However, learning techniques have yet to live up to this promise as only few methods manage to scale to high-dimensional manipulator or humanoid robots. In this talk, we investigate a general framework suitable for learning motor skills in robotics which is based on the principles behind many analytical robotics approaches. It involves generating a represen- tation of motor skills by parameterized motor primitive policies acting as building blocks of movement generation, and a learned task execution module that transforms these movements into motor commands. We dis- cuss learning on three different levels of abstraction, i.e., learning for accurate control is needed to execute, learning of motor primitives is needed to acquire simple movements, and learning of the task-dependent ”hyperparameters” of these motor primitives allows learning complex tasks. We discuss task-appropriate learning approaches for imitation learning, model learning and reinforcement learning for robots with many degrees of freedom. Empirical evaluations on a several robot systems illustrate the effectiveness and applicability to learning control on an anthropomorphic robot arm. These robot motor skills range from toy examples (e.g., pad- dling a ball, ball-in-a-cup) to playing robot table tennis against a human being and manipulation of various objects.

Biography — Jan Peters is a full professor (W3) for Intelligent Autonomous Systems at the Computer Science Department of the Technische Universitaet Darmstadt and at the same time a senior research scientist and group leader at the Max-Planck Institute for Intelligent Systems, where he heads the interdepartmental Robot Learning Group. Jan Peters has received the Dick Volz Best 2007 US PhD Thesis Runner-Up Award, the Robotics: Science & Systems - Early Career Spotlight, the INNS Young Investigator Award, and the IEEE Robotics & Automation Society’s Early Career Award. Recently, he received an ERC Starting Grant. Jan Peters has studied Computer Science, Electrical, Mechanical and Control Engi- neering at TU Munich and FernUni Hagen in Germany, at the National University of Singapore (NUS) and the University of Southern California (USC). He has received four Master’s degrees in these disciplines as well as a Computer Science PhD from USC. Jan Peters has performed research in Germany at DLR, TU Munich and the Max Planck Institute for Bi- ological Cybernetics (in addition to the institutions above), in Japan at the Advanced Telecommunication Research Center (ATR), at USC and at both NUS and Siemens Advanced Engineering in Singapore.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Estimating a Sparse Representation of Gaussian Processes Using Global Optimization and the Bayesian Information Criterion

Wilfried Wöber[1], Georg Novotny[1], Mohamed Aburaia[1], Richard Otrebski[1] and Wilfried Kubinger[1]

***Abstract*— Localization in mobile robotics is an active research area. Statistical tools such as Bayes filters are used for localization. The implementation of Gaussian processes in Bayes filters to estimate transition and measurement models were introduced recently. The non-linear and non-parametric nature of Gaussian processes leads to new possibilities in modelling systems. The high model complexity and computation expense based on the size of the dataset are shortcomings of Gaussian process Bayes filters. This work discusses our approach of a sparsing process of a dataset based on Bayesian information criterion model selection and global optimization. The developed approach combines the idea of avoiding model overfitting and Bayesian optimization to estimate a sparse representation of a Gaussian process. Based on visual odometry data of a mobile robot, the method was evaluated. The results show the operability of the system and unfold limitations of the current implementation such as random-initialization.**

I. INTRODUCTION

Bayes filters have been used frequently in mobile robotics. Different textbooks discuss the main aspects of different implementations of Bayes filters, namely Kalman filter or extended Kalman filter (EKF) [1]. Unfortunately, known restrictions limit the accuracy of Bayes filter implementations.

A Gaussian processes is a method for non-linear and non-parametric regression, which can be implemented in Bayes filters (EKF or particle filter) as a motion or measurement model [2], [3], [4]. The main benefit of a Gaussian process are estimations based on a dataset $\mathcal{D}$ including uncertainty. This leads to Bayes filter implementations, where prediction and correction are based on data [4] with minor model restrictions. The main shortcoming of Gaussian processes is the usage of the whole dataset for each estimation step. Therefore, the size of the dataset limits the processing speed.

This work tackles this problem by estimating pseudo-data for a sparse representation of a Gaussian process. This leads to the estimation of a new dataset $\mathcal{D}^*$, which consists of less data elements than the original dataset $\mathcal{D}$ without significant loss of model accuracy. This work is structured as follows: The next section discusses previous work. Section III discusses our method for optimization. Section IV evaluates our experiments. Finally, section V summarizes this work and gives an overview concerning future work.

II. PREVIOUS WORK

Bayes filters are well known methods for state estimation in mobile robotics [1, p. 23]. Doing so, $p(\vec{x}_t|\vec{x}_{1:t-1}, \vec{z}_{1:t}, \vec{u}_{1:t-1})$ must be evaluated using different approximations for motion models $p(\vec{x}_t|\vec{u}_t, \vec{x}_{t-1})$ as well as measurement models $p(\vec{z}_t|\vec{x}_t)$. This can be done using linear Gaussians in case of Kalman filter, or taylor approximation in case of EKF. To overcome approximation problems, non-parametric regression can be used to estimate models based on data. Based on that, models can be described using real system behavior. A method for such tasks is Gaussian process regression. This model is fully described using a mean and a covariance function [4], [5]:

$$GP_{\vec{\mu},\mathcal{D}}(\vec{x}_{new}) = \vec{k}^T \left[\mathbf{K} + \sigma_n^2 \mathbf{I}\right]^{-1} \vec{y} \quad (1)$$

$$GP_{\Sigma,\mathcal{D}}(\vec{x}_{new}) = k(\vec{x}_{new}, \vec{x}_{new}) - \vec{k}^T \left[\mathbf{K} + \sigma_n^2 \mathbf{I}\right]^{-1} \vec{k} \quad (2)$$

Where $GP_{\vec{\mu},\mathcal{D}}(.)$ predicts the output (mean) based on the input $\vec{x}_{new}$, the dataset $\mathcal{D}$, a kernel vector $\vec{k}$, a kernel matrix $\mathbf{K}$, the identity matrix $\mathbf{I}$ and the measurement noise σ_n^2. $GP_{\Sigma,\mathcal{D}}(.)$ predicts the inherent uncertainty using the additional scalar value $k(.)$, the kernel function. Note, that a detailed description of Gaussian processes and kernel methods can be found in [6].

The Gaussian process is based on the dataset $\mathcal{D} = \{(\vec{x}_0, y_0), ..., (\vec{x}_n, y_n)\}$, where $\vec{x} \in \mathbb{R}^{p\times 1}$ and $\vec{y} = (y_1, ..., y_n)^T$ and thus $\vec{y} \in \mathbb{R}^{n\times 1}$. Due to n examples in $\mathcal{D}$, $\mathbf{K} \in \mathbb{R}^{n\times n}$ and $\vec{k} \in \mathbb{R}^{n\times 1}$. Based on the dimensions of the Gaussian process parameters $\mathbf{K}$, $\vec{k}$ and $\vec{y}$, the size of the dataset $\mathcal{D}$ itself is critical facing real time constraints.

Gaussian process sparsing focuses on the generation of $\mathcal{D}^* = \{(\vec{x}_0^*, y_0^*), ..., (\vec{x}_m^*, y_m^*)\}$, where m is the number of examples in the new dataset $\mathcal{D}^*$ and

$$m \ll n \quad (3)$$

$$GP_{\vec{\mu},\mathcal{D}}(.) \approx GP_{\vec{\mu},\mathcal{D}^*}(.) \quad (4)$$

$$GP_{\Sigma,\mathcal{D}}(.) \approx GP_{\Sigma,\mathcal{D}^*}(.) \quad (5)$$

Recently, different approaches for Gaussian process sparsing and their applications have been discussed. In [7] a greedy sample selection is performed, where likelihood approximation is done. The subset is selected analysing the information gain. A stop criterion must be defined in terms of fixed set size or square error value. [8] generates new data points (pseudo points) to estimate $\mathcal{D}^*$ based on [7] and a maximum likelihood approach. [9] and [10] use a sparsed Gaussian process based on [8] to estimate stochastic differential equations.

Different to the previous work, the estimation of the sparse representation of a Gaussian process in this work is calculated based on the Bayesian information criterion (BIC) for pseudo input generation and global optimization for Gaussian process hyperparameter optimization.

[1] Department of Advanced Engineering Technologies, University of Applied Science Technikum Wien, Vienna, Austria, {woeber, novotny, aburaia, otrebski, kubinger}@technikum-wien.at

III. OUR APPROACH

The developed approach combines the idea of preventing model overfitting and global optimization in two stages. In the model selection stage, the sparsing of the dataset $\mathcal{D}$ using clustering and model selection is done. After that, the optimization stage optimizes a new Gaussian process to accomplish the constraints in equation 3 - 5. The remaining part of this section introduces the two stages.

A. Model Selection

The idea of sparsing in this work is based on avoiding overfitting of model selection. In this case, a finite gaussian mixture model (fGMM) was chosen to model the data. The optimal model dimension can be estimated using model selection based on the BIC [11] and a fGMM analysing $1, 2, ..., n$ mixture components. Our approach estimates the number of components using the BIC and estimates $\mathcal{D}^*$ using the expectation maximisation (EM) algorithm based fGMM fitting [12]. This is achieved using

$$p(\vec{x}|\vec{\theta}^*) = \sum_{k=1}^{m} \pi_k \mathcal{N}(\vec{x}|\vec{\mu}_k, \Sigma_k) \quad (6)$$

$$\text{where } m = \underset{j=1:n}{\operatorname{argmin}}(\text{BIC}_{\text{fGMM}}(\mathcal{D}, j)) \quad (7)$$

Where $p(\vec{x}|\vec{\theta}^*)$ describes $\mathcal{D}^*$ using a fGMM. π_j, $\vec{\mu}_j$ and Σ_j are the parameters of the j-th fGMM component, which are summarized in $\vec{\theta}^*$. m is the optimized number of pseudo-inputs based on the BIC analysis. Typically, the number of relevant samples will be smaller than the raw dataset ($m \ll n$). Note, that this assumption is based on a high number of samples. $p(\vec{x}|\vec{\theta}^*)$ is estimated using the EM algorithm. Shortcomings of this approach are discussed in chapter IV.

$BIC_{\text{fGMM}}(.)$ uses the original dataset $\mathcal{D}$ and the number of mixing components to calculate a BIC trend. This function is defined using the log-likelihood at the maximum likelihood estimation, the number of used mixture components, the sample size and the number of estimated parameters [12]. Analysing n mixing components using the BIC, the optimal model can be chosen using the minimum BIC_{fGMM} value. The sparsing is done using the mean values $\vec{\mu}_{1:m}$ of the optimized fGMM. Due to that, the sparsed dataset is $\mathcal{D}^* = \{\vec{\mu}_1, ..., \vec{\mu}_m\}$. The vectors $\vec{\mu}_{1:m}$ are called pseudo-inputs.

Note, that the discussed sparsing process tackles the optimization of the mean function of Gaussian processes. As a result of the BIC based dataset sparsing, the estimation functions are going to change. To overcome this problem, the Gaussian process hyperparameters need to be adapted. This procedure is discussed in the remaining part of this section.

B. Gaussian Process Hyperparamter Optimization

After dataset sparsing, the new Dataset $\mathcal{D}^*$ affects the mean and variance function (see equations 1 and 2). To minimize the difference between the original and sparsed Gaussian process, global Bayesian optimization was used to adapt the hyperparameters. Hyperparameter optimization is critical because of high computational effort. Simultaneously, optimization is necessary for algorithm performance. Bayesian optimization [13], [14], [15], [16] tackles this problem by reformulating the optimization to a regression problem.

Doing so, a Gaussian process again is used for this regression formulation. The main idea of Bayesian optimization is step-wise optimization based on an initialized regression model using initial samples of the optimization function. Based on those samples and a regression model, functions like the expected improvement [14], [15] evaluates the expectation and uncertainty of the regression model. The expected improvement a_{EI} is defined as [15]:

$$a_{\text{EI}}(\vec{x}|\mathcal{D}^*) = \mathbb{E}\left[\max(f^* - f(\vec{x}), 0)\right] \quad (8)$$

Where f^* is the current maximum value of the regression model and $\mathbb{E}$ is the expectation value. The function $f(.)$ returns the regression value of the regression model. Note, that different implementations extend the idea of expected improvement to control exploitation and exploration [17]. Sequential optimization is done adding an evaluation of the model to optimize at the highest a_{EI} value. In this work, we use the r^2 of the variance for model comparison. The hyperparameters of the Gaussian process are optimized in terms of optimizing the r^2.

IV. EXPERIMENTAL RESULTS

Our experiments based on measurements on a mobile robot called "Robotino"[18]. The dataset $\mathcal{D}$ is based on visual odometry calculations of five experiments. We extracted the velocity (v_x) and transition (Δx) based on those measurements. Because this paper discusses the Gaussian process sparsing, our experiments discuss the movement model sparsing in detail. Note, that the used movement model is trivial. From a machine learning perspective, the model could be represented using linear regression. Even though the model itself is simple, the Gaussian process adds uncertainty estimation, which is needed for Bayes filters.

For the analysis of our approach, we simplified the data using gathered movement information of the mobile robot. The Gaussian process based transition model was used to predict the movement of the mobile robot Δ_x along the X-axis at time t based on the velocity v_x. Additional, the implementation of our method includes data pre-processing. The data pre-processing was done using outlier elimination and data normalization. Based on our BIC based pseudo-input generation, outlier detection is critical. The used implementation uses the expectation maximization algorithm to estimate the model [12]. Due to that, implemented random cluster initialization can result in unwanted sample elimination. This would make the evaluation of $GP_{\vec{\mu},\mathcal{D}}(.)$ and $GP_{\vec{\mu},\mathcal{D}^*}(.)$ respectively $GP_{\Sigma,\mathcal{D}}(.)$ and $GP_{\Sigma,\mathcal{D}^*}(.)$ impossible.

For outlier detection, hierarchical clustering was used [19]. The software implementation is based on the hierarchical clustering functions of [20] based on euclidean distances. The visualization of the outlier detection is shown in figure 1. The algorithm classifies 26 data elements out of 4458 data elements as outliers. For further discussion, the resulting normalized 4432 data elements describe $\mathcal{D}$. The Gaussian

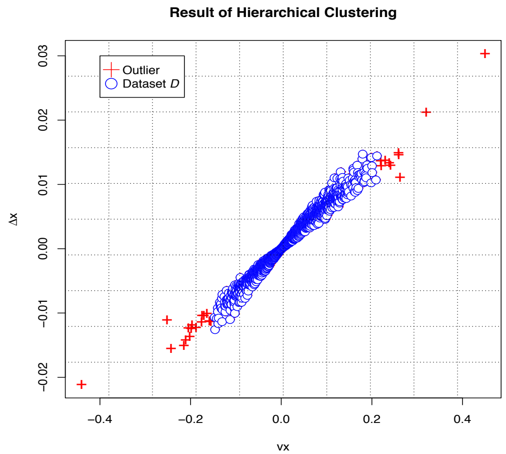


Fig. 1. Visualisation of hierarchical clustering for outlier detection

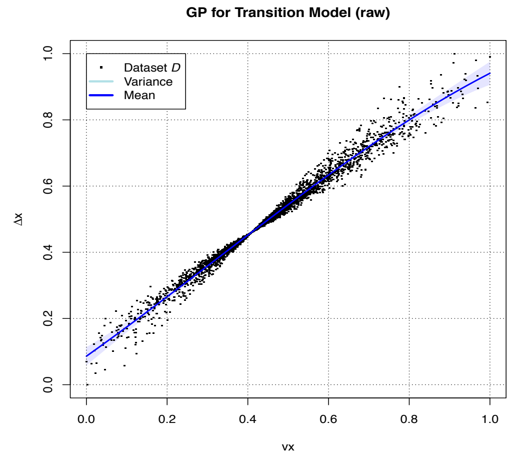


Fig. 2. Gaussian process without outliers. Note, that the data is normalized.

process based on $\mathcal{D}$ is shown in figure 2. The model sparsing was done analysing 20 to 500 pseudo-inputs using a stepsize of 50. The BIC based model selection is shown in figure 3. Note, that the implementation uses a BIC approximation which leads to a maximization instead of minimization [12]. The result of the BIC model selection is a fGMM using 170 pseudo-inputs. Those pseudo-inputs represents the dataset $\mathcal{D}^*$. Note the compression of the dataset to 170 datapoints.

Our experiments showed, that the random initialization of the fGMM clustering is critical for further optimization. The random initialization can result in a dataset $\mathcal{D}^*$, where areas with low frequency disappear. This leads to poor results of the sparsed Gaussian process. Currently, we can overcome this problem by increasing the number of datapoints in $\mathcal{D}^*$. A non-random initialization of the BIC based model selection is part of our recent research. Further, the penalty term in the function BIC_{fGMM} can be adapted for this application. The kernel used in this paper is the so-called 'rbf' kernel [4]. The hyperparameters of the kernel are the signal noise variance σ_n^2 and the smoothness factor ω [4], [6].

The behavior of the variance function is based on the hyperparameters of the Gaussian process, namely σ_n^2 and ω. Those hyperparameters were optimized using Bayesian optimization [17]. The results of the optimization are visualized in table I. The hyperparameters are optimized in 20 steps. The optimum is found at $r^2 = 0.9625$. Further, the r^2 of the Gaussian process mean values (raw and sparsed) using the optimized hyperparameters is 0.9998. Note, that due to the random initialization of the optimization algorithm, the optimization results differ. The analysis of 100 optimization procedures proves, that the exploitation/exploration tradeoff is not optimized yet and current part of further optimization. Further, due to processing limitations, 20 optimization steps and five initialization steps were used. A histogram of 100 optimization steps analysing the r^2 of $GP_{\Sigma,\mathcal{D}}(.)$ and $GP_{\Sigma,\mathcal{D}^*}(.)$ is shown in figure 4.

V. SUMMARY & OUTLOOK

We introduced a novel procedure for Gaussian process sparsing. The sparsing procedure is based on Bayesian information criterion model selection followed by hyperparameter optimization.

The model selection uses finite Gaussian mixture models to find pseudo-inputs, which represent a sparsed dataset $\mathcal{D}^*$. The hyperparameters are optimized using Bayesian optimization and focus on model difference minimization.

Our results proves that the method is applicable. Limitation, namely random initialization of model selection and optimization, are discussed. Those limitations are currently part of ongoing research. This research focuses on non-random algorithm initialization and BIC calculation adaption. Based on the results of our optimized approach, Gaussian process

TABLE I

THE OPTIMIZATION PROCEDURE IN THIS EXAMPLE.

#	σ_n^2	ω	r^2	#	σ_n^2	ω	r^2
1	3.3619	0.0171	0.7824	2	4.8541	0.0174	0.7306
3	4.0077	0.0043	0.7020	4	2.4200	0.0143	0.8156
5	0.0922	0.0086	0.9401	6	0.0050	0.0199	0.7952
7	0.0050	0.0010	0.8036	8	4.7598	0.0087	0.7048
9	0.0050	0.0121	0.7955	10	0.1486	0.0092	**0.9625**
11	0.3107	0.0041	0.9387	12	0.5753	0.0081	0.9251
13	0.4405	0.0196	0.9488	14	0.4870	0.0140	0.9407
15	0.8008	0.0190	0.9243	16	0.9876	0.0013	0.8275
17	0.3334	0.0081	0.9451	18	1.6964	0.0087	0.8340
19	2.8853	0.0093	0.7781	20	0.6128	0.0199	0.9378

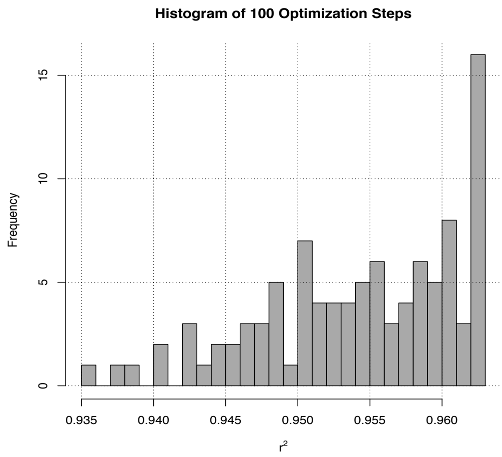


Fig. 4. Histogram of 100 optimization procedures (r^2 of $GP_{\Sigma,\mathcal{D}}(.)$ and $GP_{\Sigma,\mathcal{D}^*}(.)$).

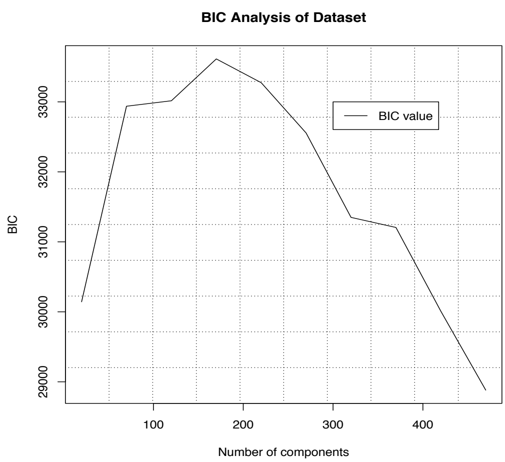


Fig. 3. Result of (approximated) BIC analysis of the transition model [12].

optimization approaches can be applied without the need of processing clouds. Currently, mobile robot localization algorithms based on sparsed Gaussian processes are implemented. This task includes the analysis of the processing workload.

Further, the expected improvement can be used to estimate the "completeness" of motion models as a preceding analysis step.

The next steps include the merging of the sparsing and optimization steps to a single optimization task. Based on the planned method extensions, non-trivial Gaussian process sparsing will be analysed. This will be used in further research areas such as example generation in object recognition.

References

[1] Thrun, S.; Burgard, W.; Fox, D., *Probabilistic Robotics*. Massachusetts Institute of Technology: MIT Press, 2006.

[2] Hartikainen, J.; Srkk, S., "Kalman filtering and smoothing solutions to temporal Gaussian process regression models," in *2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2010.

[3] Reece, S.; Roberts, S., "An introduction to Gaussian processes for the Kalman filter expert," in *2010 13th Conference on Information Fusion (FUSION)*. IEEE.

[4] Ko, J.; Fox, D., "GP-BayesFilters: Bayesian Filtering Using Gaussian Process Prediction and Observation Models," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008. IROS 2008.* Nice, France: IEEE, 2008.

[5] ——, "GP-BayesFilters: Bayesian filtering using Gaussian Process prediction and observation models," 2009, (online) https://rse-lab.cs.washington.edu/papers/gp-bayesfilter-arj-09.pdf (Last access: 18.2.2018).

[6] Bishop, C.M., *Pattern Recognition and Machine Learning*. Springer Science+Business Media, 2006.

[7] Seeger, M.; Williams, C.K.I.; Lawrence, N.D., "Fast forward selection to speed up sparse Gaussian process regression." Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics, 2003.

[8] Snelson, E.; Ghahramani, Z., "Sparse Gaussian Processes using Pseudo-inputs." Advances in Neural Information Processing Systems 18 (NIPS 2005), 2005.

[9] Garcia, C.A.; Otero, A.; Felix, P.; Presedo, J.; Marquez, D.G., "Nonparametric Estimation of Stochastic Differential Equations with Sparse Gaussian Processes." Physical Review E, 2017.

[10] Archambeau, C.; Cornford, D.; Opper, M.; Shawe-Taylor, J., "Gaussian Process Approximations of Stochastic Differential Equations." JMLR: Workshop and Conference Proceedings, 2007.

[11] Schwarz, G., "Estimating the dimension of a model." Annals of Statistics, 1978.

[12] Scrucca L., Fop M., Murphy T.B., Raftery A.E., "mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models." The R Journal, 2016, pp. 289–317.

[13] Osborne, M.A., Garnett, R., Roberts, S.J., "Gaussian processes for global optimization." 3rd International Conference on Learning and Intelligent Optimization (LION3), 2009.

[14] Bergstra, J.; Bardenet, R.; Benigo, Y.; Kegl, B., "Algorithms for hyper-parameter optimization." NIPS'11 Proceedings of the 24th International Conference on Neural Information Processing Systems, 2011.

[15] Klein, A.; Falkner, S.; Bartels, S.; Henning, P.; Hutter, F., "Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets." Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017.

[16] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 2951–2959. [Online]. Available: http://papers.nips.cc/paper/4522-practical-bayesian-optimization-of-machine-learning-algorithms.pdf

[17] Y. Yan, *rBayesianOptimization: Bayesian Optimization of Hyperparameters*, 2016, r package version 1.1.0. [Online]. Available: https://CRAN.R-project.org/package=rBayesianOptimization

[18] Festo. (2018) Robotino. [Online]. Available: http://www.festo-didactic.com/int-en/services/robotino/

[19] Liang, B., "A Hierarchical Clustering Based Global Outlier Detection Method." 5th International Conference on Bio-Inspired Computing: Theories and Applications, 2010.

[20] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Implementation of a Capacitive Proximity Sensor System for a Fully Maneuverable Modular Mobile Robot to Evade Humans

Andreas Rabl[1], Philipp Salner[1], Luis Büchi[1], Julian Wrona[1], Stephan Mühlbacher-Karrer[2] and Mathias Brandstötter[2]

***Abstract*— This paper describes an advanced approach for a dynamic collision prevention system for robots dedicated to collaborative applications in a shared human robot work environment. We developed a firmware that incorporates proximity sensor information along with a kinematic algorithm to achieve sensitive robotics for a modular mobile robot platform. The utilized sensor technology is based on capacitive sensing, capable to reliably detect humans in the vicinity of the robot platform. The kinematic algorithm is flexible in its design as it is scalable to an unlimited number of wheels and takes into account different geometric architectures such as standard and omni-directional wheels. The dynamic collision avoidance of approaching humans has been successfully demonstrated in a variety of experimental test scenarios demonstrating the capabilities of a sensitive mobile robot.**

I. INTRODUCTION

A. Motivation

The number of industrial robots in production facilities is rising steadily. The demand from the industry to have a shared work environment, where humans and robots can work together safely has increased tremendously in the last years and will become an integral part of daily work life. Further, the shortening of a product's life cycle generates the need of flexible production lines, where a sensitive and modular mobile robot platform fulfill logistics. This implies that a modular mobile robot platforms has to operate safely along with humans in a shared work environment throughout the entire time. A reliable perception system is essential to realize such a platform. The combination of kinematics of a modular mobile robot platform tightly coupled with collision avoidance technology, i.e. proximity perception sensors, are considered in this paper to safely operate a modular mobile robot platform in a shard human robot space.

B. Background

A great variety of proximity sensing technologies are available at the market such as capacitive, optical, etc. today and used in robotics. Each technology has its capabilities and comes along with benefits and limitations. Optical systems [1] have some limitation with respect to strong varying light conditions and reflections. Compared to that capacitive sensors [2] show strong non-linearities depending on the material properties and coupling to ground which can be stabilized by incorporating a proper signal processing. Thus, capacitive sensors are well known in robotics. In [3], a highly reactive collision avoidance system based on capacitive proximity sensors was evaluated. In [4] capacitive based proximity sensors were utilized on a serial manipulator to detect approaching objects in one dimension combined with a virtual compliance control of a redundant manipulator to avoid approaching objects. Further enhancements in [5] presented a contactless control of a serial manipulator based on capacitive tomographic sensors. Both works have shown that the perception system is tightly coupled to the kinematics of the robot to make them collaborative and to gain advantage of the robot's redundancy. The sensing range and characteristics of the capacitive sensors is strongly related to the geometry of the sensor front end. Investigations in [6] where done to evaluate different geometrical shapes of the sensor front.

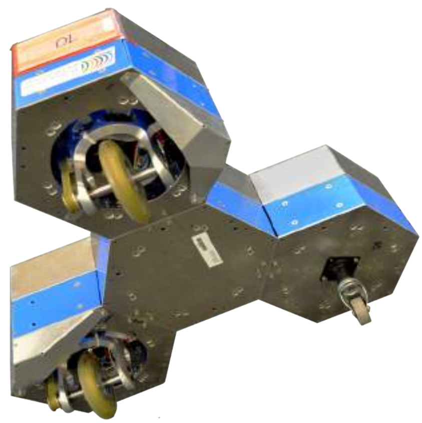


Fig. 1. Honeycomb shaped modular mobile robot platform with integrated capacitive proximity sensors.

C. Contribution

In this paper we present a fully maneuverable modular mobile robot system with integrated capacitive proximity sensors including dynamic collision prevention with humans.

[1]Andreas Rabl, Philipp Salner, Luis Büchi, Julian Wrona are with HTL Rennweg, Höhere technische Lehranstalt for Mechatronics, Rennweg 89b, 1030 Vienna, Austria andi169@gmx.at, p.salner2@gmail.com, luis.buechi@server23.cc, julianwrona1@gmail.com

[2]Stephan-Mühlbacher-Karrer and Mathias Brandstötter are with JOANNEUM RESEARCH ROBOTICS, Institute for Robotics and Mechatronics, Lakeside B08a, 9020 Klagenfurt am Wörthersee, Austria {stephan.muehlbacher-karrer, mathias.brandstoetter}@joanneum.at

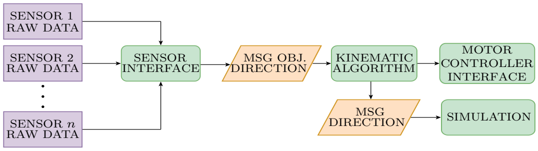


Fig. 2. Software architecture and ROS systems dependencies.

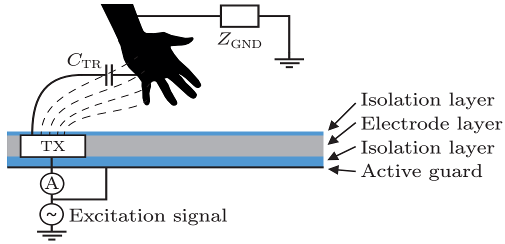


Fig. 3. Capacitive Sensing: single-ended measurement mode.

The developed advanced kinematic algorithm provides independability in terms of hardware realizations of the wheels, i.e. the modular robot platform can either consists of steered standard wheels or omni directional wheels. Furthermore, the modules of the robot can be arranged according to the needs of the application, e.g., a logistics task.

II. SYSTEM DESCRIPTION

A. Modular Wheeled Robot

The utilized modular mobile robot platform (referred to as *Wabenroboter*) consists of several hexagonal shaped submodules (referred to as hive module), each capable to be equipped with different hardware, e.g., serial manipulator. In this work two hive modules with a steered standard wheel, one hive module with a castor wheel and one hive module containing the Central Processing Unit (CPU) (Intel NUC) are utilized. The hive modules have a side length of $l_s = 150\,\mathrm{mm}$ and the main body consists of two plates stacked on top of each other, each $h_p = 90\,\mathrm{mm}$ in height. The wheel extends downwards for $h_w = 123\,\mathrm{mm}$, which results in a total height of around $h = 300\,\mathrm{mm}$. The robot geometry, as in how the hives are fixed together does not matter, for testing purposes we used the layout as shown in Fig. 1.

B. Software Architecture

The firmware consists of three main parts: The sensor signal processing module (Sensor Interface) including position estimation of an approaching human to generate a directional vector in which the robot should evade. The kinematics module (Kinematic algorithm), which determines the orientation and velocities for each wheel instantaneously. It passes the data to the module which communicates with the motor controllers (Motor Controller Interface). The overall software architecture is shown in detail in Fig. 2.

As a basis for the firmware of the robot the framework ROS (Robot Operating System) [7] is being utilized. Each part of the robots software is implemented as its own ROS package. The individual packages communicate through the ROS Publisher/Subscriber system using custom messages. To avoid communication time lags between the kinematics algorithm and the motor controller the kinematics algorithm is installed native package on the linux host. An interface class in the motor controller code enables the communication between them.

III. SENSOR TECHNOLOGY

The sensor technology in use is a capacitive proximity sensor. The measurement principle is based on the interaction of an electric field with an object approaching the sensor front end of the capacitive proximity sensor. The distortion of the electric field is caused by an object depending on its relative permittivity ϵ_r which can be measured. For proximity sensing usually the so called single-ended measurement mode is commonly utilized as illustrated in Fig. 3. In this measurement mode the capacitance between the transmitter electrode and the distant ground is determined. Therefore, an excitation signal with the frequency of $f_{ex} = 250\,\mathrm{kHz}$ is sent to each electrode in succession and the current of the displacement current is measured.

The sensor node's Printed Circuit Board (PCB) with the evaluation electronics is being supplied with $5\,\mathrm{V}$ and consists of an ultra low power wireless System on a Chip (SoC) and a 16-bit Capacitance to Digital Converter (CDC). The sensor front-end is made of a conductive copper film connected to the PCB. The measurement data is transmitted wireless with a frequency of $f_T = 2.4\,\mathrm{GHz}$ to a receiver dongle connected to the Intel NUC of the modular mobile robot platform.

The measurement characteristics of the sensor are highly dependent on the shape and size of the connected electrode's of the sensor front end which can be individually designed according to the needs of the application. In this work the size is restricted by the geometry of the hive module's side walls.

The size of the surface of the electrode, is strongly related to the maximum sensing range objects can be detected. However, increasing the size of the surface also results in the sensor being more prone to detect disturbances and noise. In Fig. 4 the shape ot the electrodes used in this work are shown.

IV. KINEMATICS

A. Kinematic System

The Wabenroboter is designed in a modular way, therefore the position and the number of wheels can change (while it is not operating). The mobile platform supports steerable standard wheels, as well as omnidirectional wheels and is configured in a way that the degree of maneuverability δ_M equals three.

The Wabenroboter is operating in a two-dimensional space so the position can be distinctly defined in ξ which holds

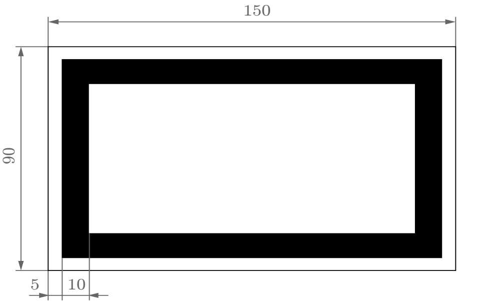


Fig. 4. Hollow shaped electrode utilized on the modular mobile robot platform.

the direction in x and y as well as the orientation angle θ. To describe the motion of the robot the values of ξ must be differentiated over the time to describe the velocity of the robot. Information on how the robot should move is received by a given trajectory which contains the velocity of the platform over time. Hence, the kinematics input is given as velocity vector $\dot{\xi}$.

$$\dot{\xi} = \begin{bmatrix} \dot{x} & \dot{y} & \dot{\theta} \end{bmatrix}^T \quad (1)$$

B. Kinematical computations

As well known from literature the kinematics of mobile robots can be modeled by using equations in the form of rolling and sliding constraints. For this work the Wabenroboter is equipped only with steerable standard wheels. These wheels are equipped with an additional vertical axis of rotation in comparison to fixed standard wheels which enables it to change β with respect to time. Hence β becomes $\beta(t)$ in the kinematic constraint equations. The vertical axis of rotation passes through the center of the wheel and the ground contact point. The rolling and sliding constraints are given for a standard steered wheel as [8]:

$$[\sin(\alpha+\beta(t)) \;\; -\cos(\alpha+\beta(t)) \;\; -l\cos(\beta(t))]\dot{\xi_R} - r\,\dot{\varphi} = 0$$
$$[\cos(\alpha+\beta(t)) \;\; \sin(\alpha+\beta(t)) \;\; l\sin(\beta(t))]\dot{\xi_R} = 0$$

In the equations above, α, l and r are geometrical values as can be seen in Fig. 5 and $\dot{\varphi}$ denotes the wheel velocity.

Much more intuitive is the geometrical view on kinematics of mobile robots. By calculating the distance of each wheel to the instantaneous center of rotation (ICR) and fulfilling the sliding constraint of the steerable standard wheel, the steering angle of each wheel is calculated. When omnidirectional wheels are used, the mobility δ_m of the robot equals three and the robot is therefore able to manipulate its position (in two-dimensional space) in every direction as well as turning around an arbitrary point. By using the rolling constraint of the equipped wheel type the rotational speed of each wheel is calculated while taking its position into account. Moreover, using the geometrical consideration the steering angle of a

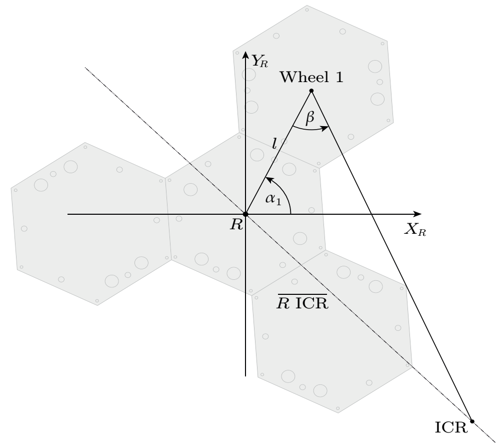


Fig. 5. Instantaneous center of rotation (ICR) and the distance to the center of the robot platform ($\overline{R\ \mathrm{ICR}}$).

standard wheel β can be calculated by

$$\beta = \arcsin\left(\frac{\overline{R\ \mathrm{ICR}}\sin(\alpha_1)}{\sqrt{l^2 + \overline{R\ \mathrm{ICR}}^2 - 2l\overline{R\ \mathrm{ICR}}\cos(\alpha_1)}}\right), \quad (2)$$

where $\overline{R\ \mathrm{ICR}}$ denotes the distance between the robots' center R and the ICR.

C. Operation

During operation (e.g., following a path) the robot has to respond to sensor input and interrupt its current task if necessary. If only omnidirectional wheels are in use, the robot can instantaneously correct its velocity vector (except of dynamical influences) and therefore react to sensor input immediately. The wheels of a mobile platform with steerable wheels must be turned correctly to allow a preferred motion. This is the reason why such drives are called pseudo-omnidirectional.

V. EXPERIMENTAL SETUP AND RESULTS

Experimental studies were done on both the robot system and the capacitive sensors. In a further step, the two systems were linked and tested together.

A. Sensor Evaluation

The characterization of the capacitive proximity sensor is performed on a linear axle for a well coupled object (similar to a human) as shown in Fig. 6. An angled profile beam is fixed on the slide of the linear axis and used to fix the electrode to avoid interferences caused by the linear axis itself. A grounded metal plate serves as the measured object. The electrode's and metal plate's surfaces are parallel during the entire test.

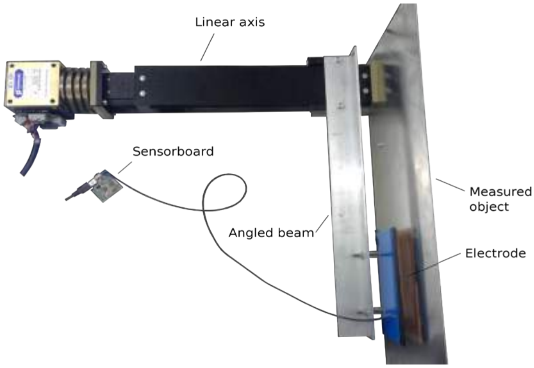


Fig. 6. Test setup to characterize the capacitive proximity sensor.

In Fig. 7 the measurement curve obtained from the test bench where an object approached the sensor front end is shown. The object is moved precisely in front of the sensor plane along $x = 0{-}200\,\mathrm{mm}$. The maximum achieved sensing range in this setup is $d_{max} = \ 60\,\mathrm{mm}$.

B. Simulation

The mobile robot platform was modeled in a simulation environment for rapid and extensive testing of the software framework. This means that even without real hardware, realistic scenarios like in the laboratory can be carried out. This was achieved by the simulation software Gazebo, which can be connected via the ROS framework, see Fig. 8.

The simulation is used during firmware development to verify the correctness of the code and visually demonstrate the entire system without using the robot. In addition to the modular mobile robot platform, the capacitive proximity sensor is also integrated into the simulation environment in order to evaluate the dynamic collision avoidance in the simulation before it is tested on the real mobile robot platform.

C. System Tests

In the experimental test setup (see Fig. 9) the modular mobile robot platform equipped with the capacitive proximity sensors drives on a predefined trajectory (sine curve) while a human approaches the robot from one side. As soon as the capacitive proximity sensor detects a human closer than $d < d_{max}$ the direction of the movement of the robot platform is changed immediately to dynamically react to the approaching human. Therefore, a contact between the human and the robot can be avoided. The modular mobile robot platform discontinues its primary task (moving on the predefined trajectory) if a human in the close surrounding of the robot is detected by the capacitive proximity sensor. If no person or object is recognized in a subsequent step, the main task is continued.

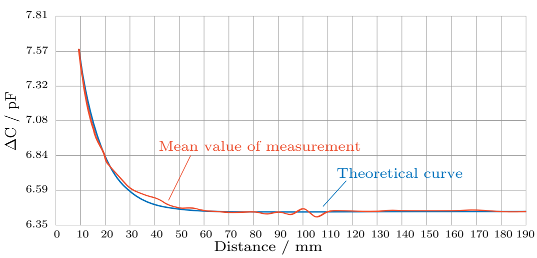
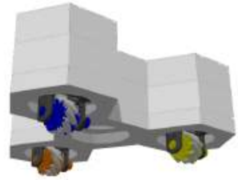


Fig. 7. Measurements of an object approaching the sensor front end of the capacitive proximity sensor.



(a) With standard wheels.



(b) With Mecanum wheels.

Fig. 8. Gazebo simulation of the Wabenroboter platform with different wheel configurations.

VI. CONCLUSIONS

In this work, a flexible firmware with capacitive proximity sensor information was developed to achieve dynamic collision avoidance for a mobile robot platform. The kinematic algorithm was developed to support various mechanical wheels and to increase the flexibility and modularity of the mobile robot platform. In addition, the integration of a capacitive proximity sensor on the modular mobile robot platform enables dynamic reaction and collision avoidance of the robot if a person approaches the robot. This enables the modular mobile robot platform to be used in a common human-robot environment. In the future a variety of electrode geometries will be evaluated to improve the sensing range of the capacitive proximity sensors.

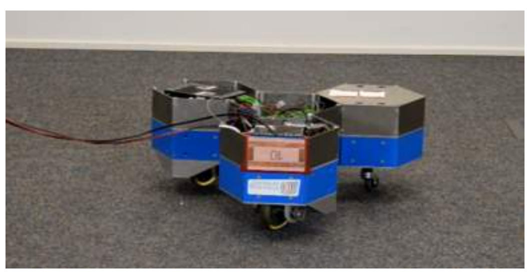


Fig. 9. Experimental test setup, where the modular mobile robot platforms executes a task and drives on a prefined trajectory (sinus curve) including dynamically collision prevention.

ACKNOWLEDGMENT

The authors would like to thank Hubert Zangl, head of the Smart System Technology Institute, Alpen-Adria-Universität Klagenfurt, supporting the work, providing the sensor hardware platform. This work was mainly funded by the Austrian Ministry for Transport, Innovation and Technology (BMVIT) within the framework of the sponsorship agreement formed for 2015-2018 under the project RedRobCo.

REFERENCES

[1] K. Terada, Y. Suzuki, H. Hasegawa, S. Sone, A. Ming, M. Ishikawa, and M. Shimojo, "Development of omni-directional and fast-responsive net-structure proximity sensor," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2011, pp. 1954–1961.

[2] S. E. Navarro, B. Hein, and H. Wörn, "Capacitive tactile proximity sensing: from signal processing to applications in manipulation and safe human-robot interaction," in *Soft Robotics*. Springer, 2015, pp. 54–65.

[3] T. Schlegl, T. Kroger, A. Gaschler, O. Khatib, and H. Zangl, "Virtual whiskers - highly responsive robot collision avoidance," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov 2013, pp. 5373–5379.

[4] M. Brandstötter, S. Mühlbacher-Karrer, D. Schett, and H. Zangl, "Virtual compliance control of a kinematically redundant serial manipulator with 9 dof," in *International Conference on Robotics in Alpe-Adria Danube Region*. Springer, 2016, pp. 38–46.

[5] S. Mühlbacher-Karrer, M. Brandstötter, D. Schett, and H. Zangl, "Contactless control of a kinematically redundant serial manipulator using tomographic sensors," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 562–569, 2017.

[6] Y. Ye, J. Deng, S. Shen, Z. Hou, and Y. Liu, "A novel method for proximity detection of moving targets using a large-scale planar capacitive sensor system," *Sensors*, vol. 16, no. 5, p. 699, 2016.

[7] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2, 2009, p. 5.

[8] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to autonomous mobile robots*. MIT press, 2011.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Analysis of Feature Tracking Methods for Vision-Based Vibration Damping of Flexible Link Robots

Florian Pucher[1], Hubert Gattringer[1] and Andreas Müller[1]

***Abstract*— Computer vision is often used in robotics where image-based feature detection is an important part. The obtained features can be exploited, e.g. for path planning, process monitoring or feedback control. In this paper the focus is on vision-based vibration damping of robots with flexible links. The measured values for control are obtained by extracting image features. The required image processing framerate depends on the link dynamics. Image processing in general is a computationally expensive task since the complexity for pixel operations is of order $O(n^2)$. Efficient algorithms for online feature tracking have to be used. In an experiment, image processing is performed on a low cost computer and results regarding the computational time are presented. The feature detection performance is validated by results of the vision-based vibration damping control.**

I. INTRODUCTION

In modern robotics applications reduction of cycle times is a critical aspect. Lightweight robots are ideal for fast operations due to their lower link inertia compared to typical industrial robots. Also power consumption is reduced. Nevertheless, the mechanical structure of lightweight robots leads to an inherent low link stiffness which causes undesireable vibrations. However, in contact with the environment increased compliance introduced by link flexibility might be even required, especially when robots are interacting with humans.

In order to damp the resulting link oscillations, addtitional sensors are required since the robotic system is underactuated. The elasticity of the links represent the unactuated degrees of freedom, see also [1].

Usually strain gauges, accelerometers or optical sensors are used for vibration control, see also [2]. In [3] strain gauges are used for curvature feedback control. Since the mounting of strain gauges is quite complex and error-prone, accelerometers are often used instead, because they are easier to apply. The acceleration measurements can be directly used for feedback control or for state estimation of flexible link robots. An example of vibration damping with accelerometers can be found in [4].

The tasks performed by robots are often monitored by an external camera system. This can be used e.g. for safety in the robot environment, process monitoring for fault detection or supervision of manipulating tasks. For guidance of the tool center point (TCP), a camera can be mounted on the robot. In this case the relative pose between the TCP and a target object can be estimated, e.g. for grasping. Since cameras are widely spread in robotic applications they can be used for vibration damping while no additional sensors are required.

Features in the camera image are used for detection of the link vibrations. The approach for vibration damping is to extend PD control of the motor angles by PD control of the feaure positions transformed into the joint space using a linearization of the image Jacobian in an operating point. The dynamics of the flexible links are modeled by concentrated elasticity in the joints (lumped element).

Different methods can be used for feature detection and tracking. The image processing rate is critical for successful vision-based vibration damping, since at least the first link eigenfrequency has to be detected. However, many algorithms have high computational costs. Therefore, in this paper some feature detection methods have been tested on a low cost computer. The first approach was the markerless estimation of the optical flow, which has already been used successfully for vibration damping of a flexible link robot in [5]. Since this approach did not work with the given setup, markers are used for feature tracking. Markers are either detected by blobs or their contours. Circular shaped markers are projected as ellipses in the image plane. With detected blobs the marker centroid is calculated. By contour detection an ellipse has to be approximated in a further processing step.

The performance of the implemented feature tracking methods is compared. Also, the quality of a feature tracking method has to be validated in combination with the vibration damping control.

II. MODELING AND CONTROL

In this section a control law for a flexible link robot using image features is presented. Also, the equations of motion used for simulation and control design are shown.

A. Dynamic Modeling

Assuming a three degrees of freedom (3-DOF) flexible link robot, the link vibrations can be modeled using a concentrated joint elasticity. This simplifying approach results in a dynamic model sufficient for the purpose of vibration damping. The equations of motion

$$\mathbf{M}_M \ddot{\mathbf{q}}_M + \boldsymbol{\tau}_f(\dot{\mathbf{q}}_M) + \boldsymbol{\tau}_A = \boldsymbol{\tau}_M \quad (1)$$

$$\mathbf{M}_A(\mathbf{q}_A)\ddot{\mathbf{q}}_A + \mathbf{g}_A(\mathbf{q}_A, \dot{\mathbf{q}}_A) = \boldsymbol{\tau}_A \quad (2)$$

$$\mathbf{K}(\mathbf{q}_M - \mathbf{q}_A) = \boldsymbol{\tau}_A \quad (3)$$

are partitioned into the dynamics of the motor angles $\mathbf{q}_M \in \mathbb{R}^3$ and the virtual link angles $\mathbf{q}_A \in \mathbb{R}^3$. The motor dynamics (1) and the link dynamics (2) are coupled via (3). The

[1]Institute of Robotics, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria, www.robotik.jku.at, {florian.pucher, hubert.gattringer, a.mueller@}jku.at

inertia matrices are $\mathbf{M}_M$ and $\mathbf{M}_A(\mathbf{q}_A)$ respectively. Motor friction is denoted by $\boldsymbol{\tau}_f(\dot{\mathbf{q}}_M)$ and the generalized motor torques are $\boldsymbol{\tau}_M$. The joint torqes $\boldsymbol{\tau}_A$ are resulting from the virtual spring stiffness matrix $\mathbf{K}$. The centrifugal and Coriolis terms, as well as link damping and gravity are combined in $\mathbf{g}_A(\mathbf{q}_A, \dot{\mathbf{q}}_A)$.

B. Camera Model

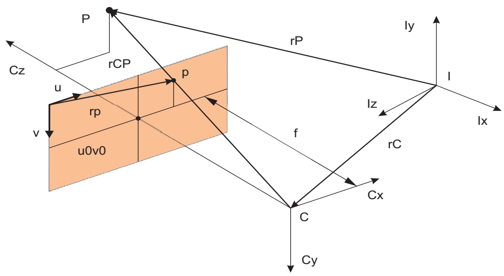


Fig. 1. Camera Model

For vision-based vibration damping a camera model, as shown in Fig. 1, is required. The perspective projection of a point P with ${}_C\mathbf{r}_{CP}^T = \begin{pmatrix} x_{CP} & y_{CP} & z_{CP} \end{pmatrix}$ onto the image plane with distance f along the optical axis ${}_Cz$ from the camera center C is

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{z_{CP}} \begin{pmatrix} f_u\, x_{CP} \\ f_v\, y_{CP} \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \tag{4}$$

The projected point is denoted by p with image coordinates $\mathbf{r}_p^T = \begin{pmatrix} u & v \end{pmatrix}$. The focal lengths f_u, f_v and the camera center (u_0, v_0) are the intrinsic camera parameters. The position vectors of C and P from the inertial point I are $\mathbf{r}_C$ and $\mathbf{r}_P$ respectively.

C. Vision-Based Vibration Damping

The vibration damping control law for flexible link robots using a camera was presented in [6]. Image feature points are transformed into the joint space and state feedback is applied. For that, a camera is mounted at the TCP (eye-in-hand). Differentiating (4) w.r.t. time leads to

$$\dot{\mathbf{r}}_p = \begin{bmatrix} \mathbf{J}_{p,v} & \mathbf{J}_{p,\omega} \end{bmatrix} \begin{pmatrix} {}_C\mathbf{v}_C \\ {}_C\boldsymbol{\omega}_{IC} \end{pmatrix} = \mathbf{J}_p\, \dot{\mathbf{z}}_C, \tag{5}$$

$$\text{with } \mathbf{J}_{p,v} = \begin{bmatrix} -\frac{f_u}{z_{CP}} & 0 & \frac{\hat{u}}{z_{CP}} \\ 0 & -\frac{f_v}{z_{CP}} & \frac{\hat{v}}{z_{CP}} \end{bmatrix}, \tag{6}$$

$$\mathbf{J}_{p,\omega} = \begin{bmatrix} \frac{u_c\, v_c}{f_v} & -\left(f_u + \frac{u_c^2}{f_u}\right) & \frac{f_u}{f_v}\, v_c \\ f_v + \frac{v_c^2}{f_v} & \frac{u_c\, v_c}{f_u} & -\frac{f_u}{f_v}\, u_c \end{bmatrix}. \tag{7}$$

Therein the image Jacobian is $\mathbf{J}_p \in \mathbb{R}^{2,6}$, the camera velocities are $\dot{\mathbf{z}}_C$, and abbreviations $u_c = u - u_0$ and $v_c = v - v_0$ are used. The image Jacobian $\mathbf{J}_p = \mathbf{J}_p(u, v, z_{CP})$ depends on the unknown distance z_{CP} of the feature point. Possible solutions for this problem are addressed in [7]. In this paper $\mathbf{J}_p(u, v, z_{CP}) \approx \hat{\mathbf{J}}_p = \mathbf{J}_p(u, v, \hat{z}_{CP})$ where $\hat{z}_{CP}$ is an approximation for z_{CP}.

The camera velocities can be expressed as $\dot{\mathbf{z}}_C = \mathbf{J}_C(\mathbf{q}_A)\,\dot{\mathbf{q}}_A$ with the geometric Jacobian $\mathbf{J}_C(\mathbf{q}_A) \in \mathbb{R}^{6,3}$ regarding the angular velocities of the links $\dot{\mathbf{q}}_A$. The unknown arm angles are replaced by the desired values $\mathbf{J}_C(\mathbf{q}_A) \approx \mathbf{J}_{C,d} = \mathbf{J}_C(\mathbf{q}_{A,d})$. The image velocities $\dot{\mathbf{r}}_p$ are given with

$$\dot{\mathbf{r}}_p = \mathbf{J}_p\, \mathbf{J}_C\, \dot{\mathbf{q}}_A. \tag{8}$$

However, for control the inverse kinematics is of interest. Since $\dot{\mathbf{r}}_p \in \mathbb{R}^2$ and $\dot{\mathbf{q}}_A \in \mathbb{R}^3$ more than one feature point is needed. For explicit calculation of the camera velocities at least three image points are required. With

$$\underbrace{\begin{pmatrix} \dot{\mathbf{r}}_{p1} \\ \dot{\mathbf{r}}_{p2} \\ \dot{\mathbf{r}}_{p3} \end{pmatrix}}_{\dot{\mathbf{r}}_F} = \underbrace{\begin{bmatrix} \mathbf{J}_{p1} \\ \mathbf{J}_{p2} \\ \mathbf{J}_{p3} \end{bmatrix}}_{\mathbf{J}_F} \dot{\mathbf{z}}_C \tag{9}$$

an inverse $\mathbf{J}_F^{-1}$ can be computed. Approximation leads to $\mathbf{J}_F^T \approx \hat{\mathbf{J}}_F^T = \begin{pmatrix} \hat{\mathbf{J}}_{p1}^T & \hat{\mathbf{J}}_{p2}^T & \hat{\mathbf{J}}_{p3}^T \end{pmatrix}$. Using a linearization of the forward kinematics at $\mathbf{q}_{A,d}$, i.e.

$$\Delta\mathbf{r}_F \approx \hat{\mathbf{J}}_F\, \mathbf{J}_{C,d}\, \Delta\mathbf{q}_A \tag{10}$$

with $\Delta\mathbf{q}_i = \mathbf{q}_{i,d} - \mathbf{q}_i\,, \quad i \in \{M, A\}$ and $\Delta\mathbf{r}_F = \mathbf{r}_{F,d} - \mathbf{r}_F$ the control law for vibration damping is

$$\begin{aligned} \boldsymbol{\tau}_M =& \mathbf{K}_{PM}\Delta\mathbf{q}_M + \mathbf{K}_{DM}\Delta\dot{\mathbf{q}}_M \\ &+ \mathbf{K}_{PA}\underbrace{\mathbf{J}_{C,d}^{+}\hat{\mathbf{J}}_F^{-1}\Delta\mathbf{r}_F}_{\approx\Delta\mathbf{q}_A} + \mathbf{K}_{DA}\underbrace{\mathbf{J}_{C,d}^{+}\hat{\mathbf{J}}_F^{-1}\Delta\dot{\mathbf{r}}_F}_{\approx\Delta\dot{\mathbf{q}}_A}. \end{aligned} \tag{11}$$

The Moore-Penrose pseudoinverse is denoted by $(\bullet)^+$. The first row in (11) is a typical PD control of motor angles used for control of robots with rigid links. The second row is used for damping of the link oscillations. Feature positions and/or velocities of feature points are required for (11).

The following sections provide a short overview of considered feature tracking methods for vibration damping. The goal is to find efficient algorithms, since the image processing rate is critical for success. Detection of the first link eigenfrequency is mandatory for this task. The methods for feature tracking are divided into markerless and marker-based techniques.

III. Markerless Feature Tracking

Typical features in an image are edges, corners and blobs. Without addtional information about the features or the camera scene, corners are best suited for tracking. Blob detection can be especially useful if markers of known shape, size or color are used. The detection of edges can be used for finding contours of objects. Blob and edge detection are subjects of section IV concerning marker-based feature tracking methods.

A. Corner Detection

Corners can be found with the Shi-Tomasi corner detector or the Harris corner detector [8]. In [9] an improvement of the selection criteria for corners compared to the Harris corner detector is presented. For this reason both algorithms require approximately the same amount of computational time.

B. Optical Flow

The optical flow is a vector field describing the relative displacement of pixels between two consecutive frames of a video. The calculated pixel velocities can be used in (11). The differential methods for estimation of the optical flow are based on the assumption that the illumination $I(u, v, t)$ between two subsequent frames is constant. The equation

$$\frac{dI}{dt} = \frac{\partial I}{\partial u}\dot{u} + \frac{\partial I}{\partial v}\dot{v} + \frac{\partial I}{\partial t} = 0 \tag{12}$$

is the basis for calculation. The optical flow can be computed using, e.g. the Horn-Schunk method [10] or the Lucas-Kanade method [11].

Dense algorithms compute the optical flow for each pixel, whereas the sparse techniques rely on features. Only sparse algorithms, as the pyramidal implementation of the Lucas-Kanade method [12], are considered here.

Since image corners can vanish over time, in each image a new set of corner features is detected and tracked in the consecutive image. This means the method using optical flow only supplies feature velocities but no feature positions. The vibration damping is achieved solely by feedback of image velocities, i.e. by setting $\mathbf{K}_{PA} = \mathbf{0}$ in (11).

IV. Marker-Based Feature Tracking

The use of objects (markers) of known size, shape and color can greatly reduce the processing time of feature tracking. Since the main goal is the verification of the control law (11) for vibration damping, the image environment is constructed to have only few textures. This makes it easier to detect the markers and reduces computational effort. In this paper three black circular markers on a light gray background are used. Due to projection into the image plane elliptic markers have to be assumed. These markers can be detected by either the blob regions or the boundaries of the regions, i.e. the contours.

A. Region of Interest

A method for vastly reducing the computational effort is the use of small image areas, the regions of interest (ROI), where the image processing is performed. The size of the ROI is chosen by using the knowledge of the marker size in the image and the expected displacement of the marker. The ROI are centered around the feature position of the preceding image.

B. Blob Detection

A basic and fast method for marker detection is the generation of a binary image using thresholding. This seperates markers from the background. The detected blobs by using thresholding can be either used directly for estimation of marker properties or further processed, e.g. by extraction of the contour.

For conversion of a gray scale image with pixel intensity $I(u, v)$ into a binary image with $I_b(u, v)$ a decision based on a threshold value I_{Th} is used. If the gray level is greater than the threshold, the resulting pixel is white. If not, it is a black pixel, i.e.

$$I_b(u, v) = \begin{cases} 1 & \text{if} \quad I(u, v) > I_{Th} \\ 0 & \text{if} \quad I(u, v) \leq I_{Th} \end{cases} . \tag{13}$$

For varying illumation across different image regions an adaptive threshold can be used. Constant threshold is more efficient here, because for each ROI a different value can be used. In Fig. 2 on the left hand side a gray scale image is shown and on the right hand side is the resulting binary image for a constant threshold value.

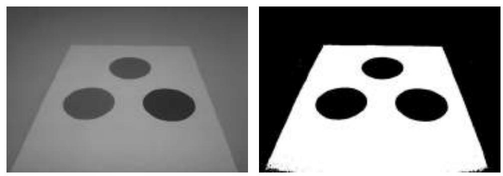


Fig. 2. Blob detection by thresholding

C. Contour Detection

Detection of contours can be done by finding the borders of blobs or by edge detection.

1) Border Following: In a binary image there are regions of black pixels adjacent to white pixels. The contours are the connected components found by checking the pixel neighbourhood. A border following algorithm is presented in [13]. In Fig. 3 the found contours are shown for the full image only for demonstration purposes. For efficient calculation the regions of interest are used.

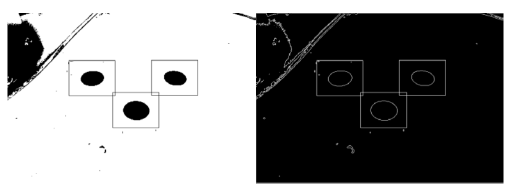


Fig. 3. Binary image and contours

2) Edge Detection and Linking: Edges are typically detected in a gray scale image using the Canny edge detector, see [14]. The edges, however, are not connected in general. If edges are used, they need to be connected to obtain the full contours of an object.

D. Marker Position Estimation Methods

Having the blobs or contours detected, the next processing step is the extraction of the marker positions. The blob centroid is an appropriate candidate for the marker position. In case of detected ellipse contours, the ellipse parameters have to be estimated. Possible methods are based on least-squares techniques or Hough transform. The center of the ellipse is the wanted marker position.

1) Statistical Moments: Shape information of detected blobs can be extracted by the use of statistical moments. Here a binary image with pixels $I_b(u, v) \in \{0, 1\}$ is assumed although the concept of statistical moments is more general and can be also used for gray scale images. The statistical moment of order $p + q$ is defined as

$$m_{pq} = \sum_{u,v \in \mathcal{I}} u^p \, v^q \, I_b(u, v) \tag{14}$$

within a region $\mathcal{I}$. The location of a marker is required for vibration damping. This can be e.g. the centroid

$$\begin{pmatrix} \overline{u} & \overline{v} \end{pmatrix} = \frac{1}{m_{00}} \begin{pmatrix} m_{10} & m_{01} \end{pmatrix} \tag{15}$$

of the marker. Thresholding and calculation of the moments can be efficiently done within only one loop over the pixels. The decision if a blob is the wanted marker, can be based on the area

$$m_{00} = \sum_{u,v \in \mathcal{I}} u \, v \, I_b(u, v) \tag{16}$$

of the blob and the previous marker location.

2) Ellipse Approximation Using Least-Squares: The equation for a general ellipse in image coordinates (u, v) is

$$a_{uu} \, u^2 + a_{uv} \, u \, v + a_{vv} \, v^2 + a_u \, u + a_v \, v + a_0 = 0 \tag{17}$$

with the parameters $\mathbf{a}^T = \begin{pmatrix} a_{uu} & a_{uv} & a_{vv} & a_u & a_v & a_0 \end{pmatrix}$. With a given set of contour points an ellipse is approximated. An algorithm for least-squares fitting is presented in [15]. The method is based on eigenvector calculation. The center of the ellipse is the feature position used for vibration damping.

The gray scale image with the detected ellipse contours in the regions of interest is shown in Fig. 4. The centers of the ellipses are also drawn in the figure. The least-squares fitting can be also successful if some parts of the contour are missing.

3) Ellipse Extraction Using Hough Transform: Using the Hough transform geometric objects like lines, circles or ellipses can be found in a contour image. Based on the equation of the corresponding geometric object the parameter space is quantized. For each set of parameters an accumulator is increased if the equation is fulfilled for a pixel. The computational effort increases with a high dimensional parameter space, since the method is like a brute-force algorithm. An ellipse has a five dimensional parameter space, therefore the hough transform is quite computationally expensive.

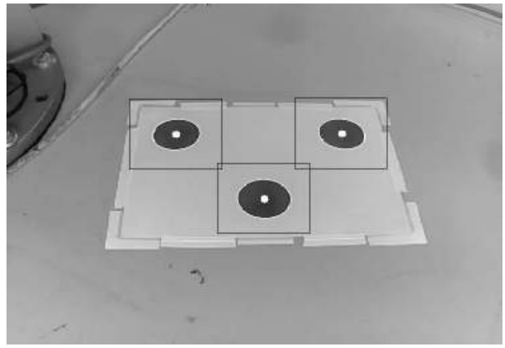


Fig. 4. Gray scale image including ROI and detected ellipses

E. Marker Tracking

If more than one ellipse is detected within a region of interest, the one with the smallest euclidian distance from the previous ellipse is chosen. In this case the feature tracking does not require any additional image processing operations.

V. EXPERIMENTAL SETUP AND RESULTS

For the experiment a Raspberry Camera is mounted at the TCP of the 3-DOF flexible link robot ELLA (Elastic Laboratory Robot), developed at the Institute of Robotics at the Johannes Kepler University Linz. The experimental setup is schematically shown in Fig. 5. Three black circular markers are placed in front of the robot with a distance of ca. 0.06 meters from the camera. The first link eigenfrequency lies within the range of 4 to 5 Hertz. The maximum frame rate of the camera is 90 frames per second. Image processing is performed on a *Raspberry Pi 2* for gray scale images with a resolution of 640×480 pixels. The size of the ROI is 120×100 pixels for each of the three regions.

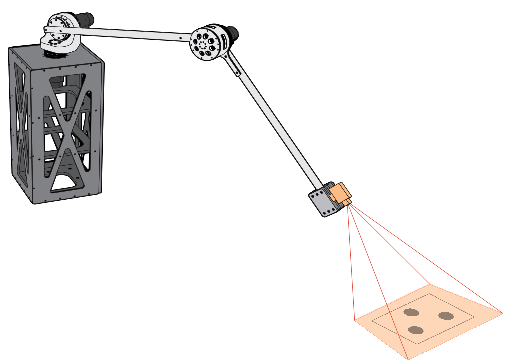


Fig. 5. Elastic robot with camera-setup

A. Feature Tracking Performance

With the Shi-Tomasi corner detector and the pyramidal implementation of the Lucas-Kanade feature tracker an image processing rate of 25 frames per second is achieved by using a full image. This method was tested in an environment with more image textures than in the marker-based ones. Compared to the first link eigenfrequency the computational time is too high for vibration damping with the given experimental setup.

Hough transform was only tested for circles, which have a three dimensional parameter space, and was omitted immediately, since the performance was insufficient and the computational effort for ellipses is higher than for circles.

With the method using (15) for blob centroid calculation in a thresholded image, 90 frames per second are obtained. The marker detection algorithm with the steps

1) Threshold
2) Contour Detection by Border Following
3) Least-Squares Ellipse Fitting

also reaches an image processing performance of 90 frames per second. The last described algorithm was used in the validation experiment.

B. Experimental Results

In Fig. 6 and Fig. 7 the vision-based vibration damping is compared to the undamped case, i.e. pure PD control of the motor angles. The excitation signal is a step disturbance of $1\,\mathrm{N\,m}$ in the motor torques. The resulting link vibrations are successfully damped. In this experiment no feature velocities were used in the control law (11), i.e. $\mathbf{K}_{DA} = \mathbf{0}$.

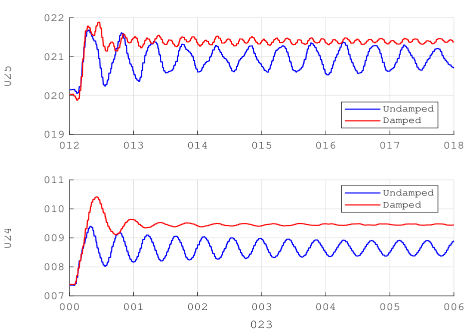


Fig. 6. Image coordinates (u_1, v_1) of a marker

In the horizontal oscillations, i.e. Fig. 6 (top) and Fig. 7 (top), it is obvious that vibrations with higher frequencies are difficult to damp using vision-based control.

VI. CONCLUSIONS

In vision-based control feature tracking is a challenging task. Especially for robots with flexible links, where the objective is vibration damping, fast motion detection is required. Therefore, in this paper different feature tracking

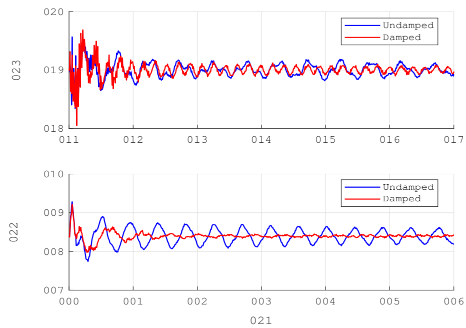


Fig. 7. TCP acceleration response to a step disturbance

methods were analyzed regarding computational efficiency. Marker-based techniques have the advantage of greatly reducing the size of processing data with few image processing steps due to a-priori knowledge of the objects. With two methods the maximum possible image processing rate for the given camera setup was achieved. The efficiency of the feature tracking was validated by a vibration damping experiment.

Investigations on calculating the optical flow within ROI will be done in the future. Also a comparison between marker position extraction from contours and blobs regarding robustness is of interest. Furthermore, the use of edges for contour detection has to be implemented and compared to border following.

ACKNOWLEDGMENT

This work has been supported by the Austrian COMET-K2 program of the Linz Center of Mechatronics (LCM), and was funded by the Austrian federal government and the federal state of Upper Austria.

REFERENCES

[1] R. Seifried, *Dynamics of Underactuated Multibody Systems: Modeling, Control and Optimal Design*, ser. Solid Mechanics and Its Applications. Springer International Publishing, 2013.

[2] B. Siciliano and O. Khatib, *Springer Handbook of Robotics*, ser. Springer Handbooks. Springer International Publishing, 2016.

[3] W. Höbarth, H. Gattringer, and H. Bremer, "Modelling and control of an articulated robot with flexible links/joints," in *Proceedings of the 9th International Conference on Motion and Vibration Control*, 2008, p. 10.

[4] P. Staufer and H. Gattringer, "State estimation on flexible robots using accelerometers and angular rate sensors," *Mechatronics*, vol. 22, no. 8, pp. 1043 – 1049, 2012.

[5] J. Malzahn, A. S. Phung, R. Franke, F. Hoffmann, and T. Bertram, "Markerless visual vibration damping of a 3-dof flexible link robot arm," in *ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics)*, 2010, pp. 1–8.

[6] P. Florian, G. Hubert, and M. Andreas, "Vibration damping of flexible link robots using an eye-in-hand camera," *PAMM*, vol. 17, no. 1, pp. 163–164, 2018.

[7] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*, ser. Advanced Textbooks in Control and Signal Processing. Springer London, 2010.

[8] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey vision conference*, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.
[9] J. Shi and C. Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
[10] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1, pp. 185 – 203, 1981.
[11] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, ser. IJCAI'81. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1981, pp. 674–679.
[12] J. yves Bouguet, "Pyramidal implementation of the lucas kanade feature tracker," Intel Corporation, Microprocessor Research Labs, Tech. Rep., 2000.
[13] S. Suzuki and K. be, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32 – 46, 1985.
[14] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
[15] A. W. Fitzgibbon and R. B. Fisher, "A buyer's guide to conic fitting," in *Proceedings of the 6th British Conference on Machine Vision (Vol. 2)*, ser. BMVC '95. Surrey, UK, UK: BMVA Press, 1995, pp. 513–522.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
 ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Autonomous Extrinsic Calibration of a Depth Sensing Camera on Mobile Robots

Farhoud Malekghasemi[1], Georg Halmetschlager-Funek[1] and Markus Vincze[1]

***Abstract*— This work presents a fast and autonomous system to find the rigid transformation between the RGB-D camera and a local reference frame on a mobile robot. The major advantages of the method over the conventional methods of calibration is that there is no need for a special setup or any known object in the scene and its speed. This is achieved by taking advantage of robot's motion combined with camera tracking method. We show that two circular motion and one plane detection are sufficient to autonomously calibrate the robot in the different environments with some minimal texture. The presented method is evaluated with both, computer simulation and in real-life scenarios.**

I. INTRODUCTION

Depth sensing cameras like Microsoft's Kinect, also known as RGB-D cameras, provide robots with three-dimensional information (3D) of its environment by using structured infrared light (cf. Fig. 1). Therefore, they have become very popular especially in the branch of mobile robotics because the depth perception is necessary for a successful obstacle avoidance, SLAM, object recognition, segmentation, 3D reconstruction and camera tracking [1]–[4]. 3D camera collects information from its own perspective (in camera coordinates), which is then transformed to a global coordinate system in order to relate the other parts of the robot to achieve a required task. The transformation is only possible when the relationship between the camera and other parts of the robot are known, therefore most methods require prior knowledge of accurate measurement. The parameters which are used to describe this relationship are called extrinsic parameters of a camera.

In practice, extrinsic parameters of a camera are not always constant and could vary during the time in multiple cases such as wear and tear in robot parts, collision accidents, changing camera's mounting place on robot's body by the user to adopt different environments. All of these displacements violate the prior assumption of known transformations. Thus, recalibration of the camera is necessary. But the process of recalibration is challenging and time-consuming. The state-of-the-art methods of extrinsic calibration are to measure distances directly or using reference objects on the scene with precalibrated position and orientation [5] which are respectively, not accurate and needs a long procedure. These methods are not easily repeatable without an expert

[1]All authors are with the V4RLab, ACIN, Faculty of Electrical Engineering, TU Wien, 1040 Vienna, Austria. {farhoud.malekghasemi, georg.halmetschlager, markus.vincze}@tuwien.ac.at

This work is supported by the European Commission through the Horizon 2020 Programme (H2020-ICT-2014-1, Grant agreement no: 645376).

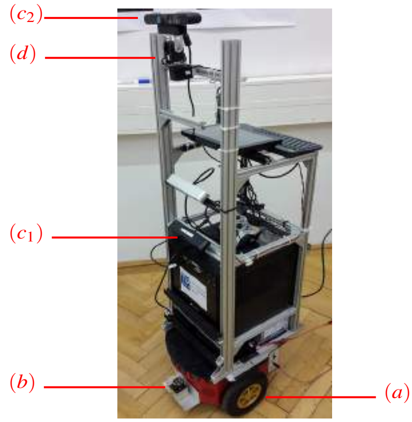


Fig. 1. V4core is a mobile robot system for research and development made by Vision4Robotics group in Vienna University of Technology based on a Pioneer P3-DX platform from MobileRobots company. It is equipped with (a) two differential-drive wheels, (b) scanning laser in front, $(c_{1,2})$ Two Asus Xtion RGBD cameras which are mounted at different heights and (d) a pan-tilt unit for top camera.

in the loop, e.g., for filling the images into the the system with showing a calibration template to the cameras.

To overcome these problems we contribute a fast and autonomous method to find the rigid transformation between the RGB-D camera and a reference frame on a mobile robot by taking advantage of its motion combined with visual motions estimation of the camera. We assume the robot's working area is a flat floor with some minimal texture, on which it can freely move around and is observed by the camera.

II. RELATED WORK

Most of related works focus on the calibration of intrinsic camera parameter. There exist several methods that use the motion of a camera to calibrate the intrinsic parameters, such as: [6]–[8].

Calibrations without using any specialized reference objects or patterns have also been studied in [9]–[12]. Carrera et al. [10] calibrated a fixed multi RGB camera rig by detecting invariant SURF feature correspondences across different views. In Miller et al. [11] work, the extrinsic parameters were estimated for calibrating the relative pose and time offsets of a pair of depth sensors based on point correspondences established from the unstructured motion of

objects in the scene. Pathirana et al. [12] proposed a method to calibrate multiple cameras based on users joint positions.

Furthermore, similar techniques are proposed for calibrating 2D and 3D LIDAR sensors mounted on a moving platform in [13]–[15].

III. METHOD

The proposed method estimates the camera pose in 3D space by driving the robot in predetermined paths which provide us the required camera trajectories for calculations. The pose of the camera in 3D space is described by translation and rotation with respect to the robot's base coordinate system. Therefore, there are six parameters (6 DoF) to be determined:

- Three translation distances: X, Y, Z
- Three rotation angles: Roll(ϕ), Pitch(θ), Yaw(ψ)

We divide these six parameters in three categories as following:

1) Z, ϕ, θ
2) X, Y
3) ψ

to design a special step for separately calculating parameters of each category. These steps have been named respectively as ground plane detection, two-rotation drive and straight-forward drive. Fig. 2 shows an overview of three main steps and relation between sub steps in the algorithm.

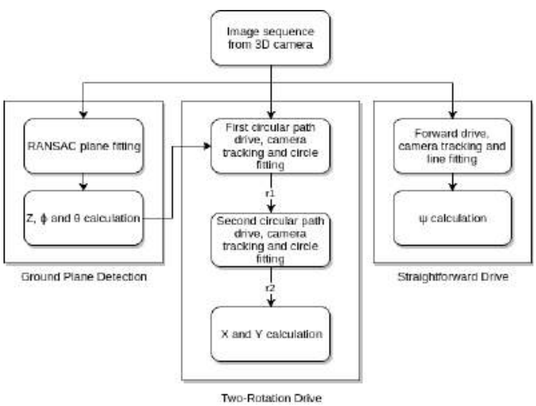


Fig. 2. Overview of different steps in the approach.

A. Camera Tracker

A camera tracking algorithm provides us the trajectory of the camera in 3D space. This trajectory is determined using visual motion estimation with the suggested method in [16]. The method combines two approaches:

- Feature-based method using a pyramidal implementation of the KLT-tracker [17].
- A keyframe-based refinement step.

It detects FAST-keypoints [18] first to initialize a keyframe and assign them to the corresponding 3D locations. Then it tracks, frame by frame, the keypoints using pyramidal KLT-tracker, which allows tracking large camera motions. Finally, it uses RANSAC to robustly estimate the rigid transformation (the camera pose) from the corresponding depth information of the organized RGB-D frames. Additionally, it applies a keyframe-based refinement step by projecting patches to the current frame to account for the accumulated drift for individual point correspondences and optimizing their locations. The algorithm produces as output a set of keyframes $K = \{K^1, ..., K^n\}$ and a set of transformations $T = \{T^1, ..., T^n\}$ for camera pose adjusting the corresponding keyframes to the reference frame which is defined by the first camera frame or by the user.

B. Ground Plane Detection

We assumed that the robot is working on a flat floor. Detecting this one plane in the camera's field of view is enough to calculate the parameters of the first category. The segmentation algorithm finds all the points within a point cloud that supports a plane model using the random sample consensus (RANSAC) [19] as a robust estimator of choice. A threshold for distance determines how close a point must be to the model in order to be considered as an inlier. Finally the contents of the inlier set, are used to estimate coefficients of plane's equation in 3D space:

$$n_x x + n_y y + n_z z + d = 0 \tag{1}$$

wherein d represents the distance between the plane and the camera, which is equivalent to the distance of the camera from the ground (height of the camera), also known as Z parameter in this category. The vector $\mathbf{n}^T = [n_x, n_y, n_z]$ represents the normalized normal vector of the plane which is perpendicular to the surface. The formed angles between this normal vector and the camera's coordinate system provide the roll and pitch angles in this category, which can be simply calculated using trigonometry as illustrated in Fig. 3.

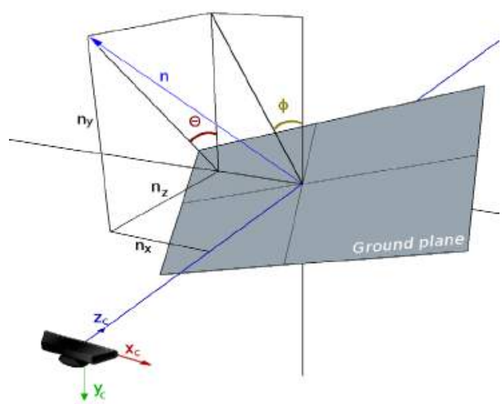


Fig. 3. Roll(ϕ), Pitch(θ) with respect to detected ground plane in camera coordinate system.

The pitch angle of camera θ is equal to the angle between the normal vector of the ground plane $\mathbf{n}$ and the $x_c - y_c$ plane of camera coordinate system, so it is calculated with:

$$\theta = \arctan\left(\frac{n_z}{n_y}\right). \tag{2}$$

The roll angle of camera ϕ is equal to the angle between the normal vector of the ground plane $\mathbf{n}$ and the $y_c - z_c$ plane in the camera coordinate system, so it is calculated with:

$$\phi = \arctan\left(\frac{n_x}{n_y}\right). \tag{3}$$

C. Two-Rotation Drive

This method is used to calculate the second category parameters, including X and Y distances of the camera in the robot base coordinate system. In case of mobile robots, the base coordinate system is usually chosen to be at the robot's center of mass. In this step, the robot rotates along two circular paths with different radiuses. If the trajectory of the camera is determined during these rotations, then X and Y distances can be calculated using simple geometry.

The camera tracker provides these transformations from camera perspective in the base frame of the robot, which we selected as the reference frame. Before any calculation is started the camera trajectory should be transformed to compensate for roll and pitch angels that have been found in the previous section since the camera coordinate and robot base coordinate systems are rotated respectively.

When the robot drives two times with circular path, the camera also has circular movement trajectories with respect to the center of rotation. Radiuses for these two drive can be chosen arbitrary but they should not be equal distances. In order to reduce path execution error and noise production during the drive, rotation radius is selected to be zero for the first drive and half of the distance between two wheels for the second drive. In the first rotation, the robot rotates exactly around itself on a spot and the center of rotation will be equal to the base origin. Then in the second rotation, it will rotate exactly around one of the wheels which keeps this wheel's motor off. Fig. 4 shows a robot from top view with two differential-drive wheels in the base coordinate system and two camera trajectories that would be taken during the first and second circular path drive by the camera mounted on it.

The camera trajectory is a set of points T in 3D space in camera coordinate system which has rotation with respect to base coordinate system. Therefore before any further more calculation this trajectory should be rotated using roll and pitch angels determined in previous section. The yaw angle does not effect calculation because we are only interested in magnitude of circle's radius and since the camera height is fixed trajectory's data in the z direction is also irrelevant here. r_1 and r_2 radiuses can be calculated by applying 2D circle fit algorithm on this set of points.

The algorithm is an implementation of direct least squares fitting a circle to 2D points in [20]. The goal is to fit a set of points with a circle equation:

$$(x-a)^2 + (y-b)^2 = r^2 \tag{4}$$

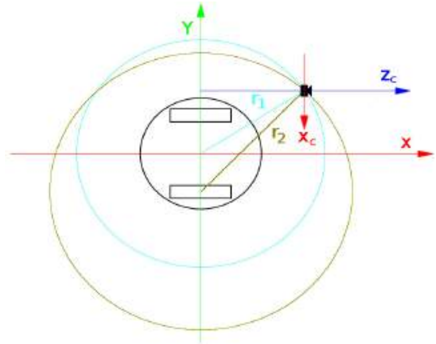


Fig. 4. Top view of a robot with two differential-drive wheels in the base coordinate system and a 3D camera on board for Two-Rotation drive.

where $[a,b]^T$ is the circle center and r is the circle radius. The error function to be minimized for n points in the set is:

$$E = \sum_{i=1}^{n} (L_i - r)^2 \tag{5}$$

where $L_i = \sqrt{(x_i - a)^2 + (y_i - b)^2}$. Setting to zero of partial derivatives of equation 5 with respect to a, b and r leads to:

$$r = \frac{1}{m} \sum_{i=1}^{n} L_i \tag{6}$$

$$a = \frac{1}{m} \sum_{i=1}^{n} x_i + r \frac{1}{m} \sum_{i=1}^{n} \frac{\partial L_i}{\partial a} \tag{7}$$

$$b = \frac{1}{m} \sum_{i=1}^{n} y_i + r \frac{1}{m} \sum_{i=1}^{n} \frac{\partial L_i}{\partial b}. \tag{8}$$

These equations can be solved using fixed-point iteration to obtain radius r and center of the circle.

After calculation of r_1 and r_2, the two camera trajectories equations in the $x - y$ plane can be written as below:

$$x^2 + (y - r_F)^2 = r_1^2 \tag{9}$$

$$x^2 + (y - r_S)^2 = r_2^2 \tag{10}$$

with r_F and r_S are the first and second drive radiuses. X and Y distances are calculated by solving two circles equations for an intersection point as the flowing:

$$Y = \frac{r_1^2 - r_2^2 + r_S^2 - r_F^2}{2(r_S - r_F)} \tag{11}$$

$$X = \sqrt{r_1^2 - (Y - r_F)^2}. \tag{12}$$

We assume that the camera is looking forward on the robot, therefore the calculated negative value for X will be discarded.

D. Straightforward Drive

This method is used to calculate the yaw angle of the camera in the last category. For this calculation, the robot starts driving straightforward for a short distance while the camera tracker is providing the camera trajectory. As shown in Fig. 5, considering camera trajectory with respect to the camera coordinate system forms a line in the $z_c - x_c$ plane. Calculation of the angle between z_c axis of the camera and this line yields to the yaw angle. In order to get the slope of the trajectory, a line fit algorithm can be applied to the set of camera trajectory points.

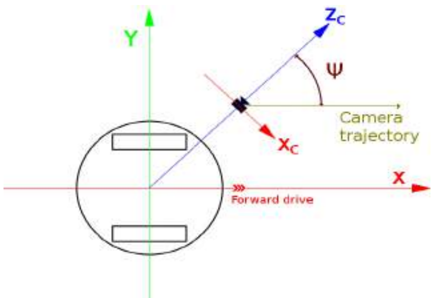


Fig. 5. Top view of a robot with two differential-drive wheels in the base coordinate system and a 3D camera on board for straightforward drive.

The line fit algorithm is an implementation of linear fitting of 2D points in [20]. The goal is to fit a set of points with a line equation:

$$y = Ax + B. \tag{13}$$

The error function to be minimized is the sum of the squared errors between the y values and the line values (only in y-direction).

$$E = \sum_{i=1}^{n} [(Ax_i + B) - y_i]^2 \tag{14}$$

Setting gradient of equation 14 to zero leads to a system of two linear equations:

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix} \tag{15}$$

which can be solved to obtain A and B.

$$A = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i} \tag{16}$$

$$B = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i} \tag{17}$$

Finally, the yaw angle is calculated with:

$$\psi = \arctan(A). \tag{18}$$

IV. IMPLEMENTATION

The V4core (Fig. 1) mobile robot platform is used for implementation of the presented method and obtaining data in real-life scenarios. We used the planar segmentation algorithm from Point Cloud Library (PCL) [21] in the ground plane detection step of the method, and the V4R[1]-library's [22] camera tracker to obtain camera movements trajectory in two other steps.

V. EXPERIMENT

Three cameras are mounted at different heights and angles on the robot looking to the floor in front of it. Ground truth for X, Y and Z lengths of these cameras are measured manually using tape measure and laser measuring tool and the ground truth angels are measured by cameras looking at a fiducial marker (similar to QR code) that is fixed in the environment of the robot. Afterwards this data is used to build a simulation model of the robot in GAZEBO simulation environment in which the proof-of-concept tests are conducted. Since the camera tracker is not able to work properly in simulation environment a piece of software is developed to calculate camera trajectory based on the tf tree of the robot and simulate the camera tracker. The outcome proved functionality of the method under ideal circumstances of simulation compare to ground truth data (cf. Fig. 7, 8, 9).

For real-life scenarios, two kinds of experiments are conducted in areas with mosaic and wooden floor structure (Fig. 6) several times for each of the three cameras. The results are gathered from the bottom camera at 0.44 m, middle camera at 0.75 m and top camera at 1.33 m from the floor. The whole process of calibration for each camera took under three minutes. Fig. 7, 8, 9 demonstrate estimation error and standard error of it with respect to ground truth for each camera in both areas. Fig. 10, 11 illustrate comparison between the calculated ground truth camera trajectories and the obtained results from real camera tracker on both areas. These measurements are recorded during a circular path drive of the robot.



Fig. 6. Mosaic and wooden floor structure.

Bottom camera provides very good tracking results in both areas (cf. Fig. 10, 11). Because the camera tracker trajectories are almost in a perfect circle shape and fitted circles to them are very close to the ground truth trajectories. This type of good matching results in accurate estimation of X and Y lengths with less than 10 mm error (Fig. 7) in both areas. But the pose estimation error increased for the middle camera and

[1]Vision for Robotics group. ACIN, TU Vienna

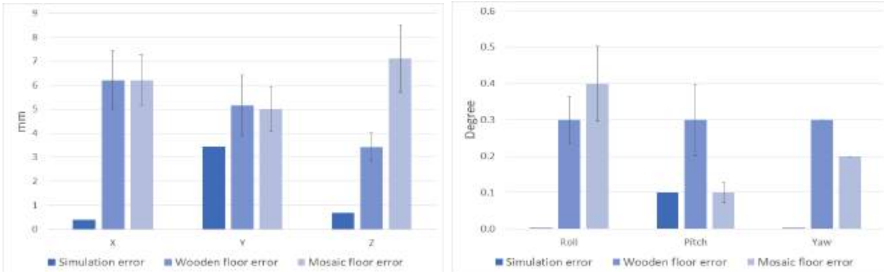


Fig. 7. Bottom camera pose error (X = 0.28 m, Y = -0.15 m, Z = 0.44 m, Roll = 1.1°, Pitch = 26.9°, Yaw = -29.8°).

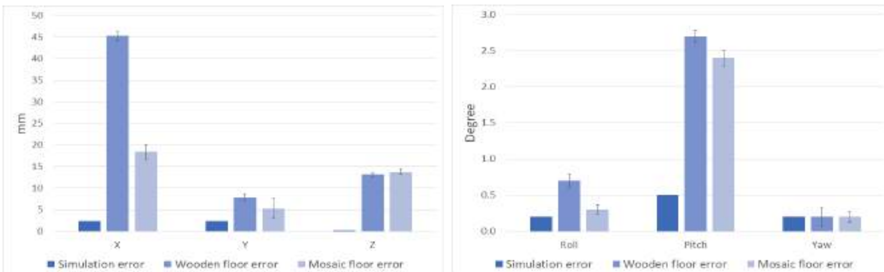


Fig. 8. Middle camera pose error (X = 0.115 m, Y = 0.02 m, Z = 0.755 m, Roll = 0.1°, Pitch = 34.4°, Yaw = 0.1°).

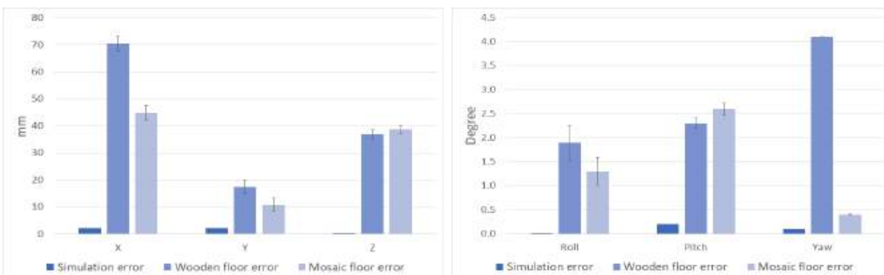


Fig. 9. Top camera pose error (X = 0.173 m, Y = 0.02 m, Z = 1.333 m, Roll = 0°, Pitch = 51.6°, Yaw = 0°).

it increased dramatically for the top one (Fig. 8, 9), because the camera tracking error got bigger as the height of cameras from the floor increased (Fig. 10, 11). The better performance of the camera tracker on the mosaic floor area compared to the wooden one is also noticeable (cf. Fig. 10, 11) which leads to less estimation error for mosaic area. The reason behind this is that the mosaic floor has a lot more traceable features (texture) in its pattern than the wooden floor.

Another noticeable point in both scenarios is the increment of error in estimation of the Z length and the angels, when the camera is getting far from the floor. This result was expected because it is already shown in previous works when the distance between the 3D camera and the planar surface increases the depth accuracy of the sensor decreases [23].

VI. CONCLUSION

This paper focused on presenting a novel autonomous and fast method for extrinsic calibration of a 3D camera on board of a mobile robot without any need for artificial targets, using camera motion estimation and robot's mobility. The simulation results proved the concept and the real-life scenarios also demonstrated that, with consideration of the accuracy range of the depth sensor and sufficient texture of the robot's working environment, it can provide good results in term of accuracy for practical cases. It is known from stereo systems that the floor always contains some texture or stains which are sufficient to be tracked contrary to walls, that might be really textureless. The method has a significant

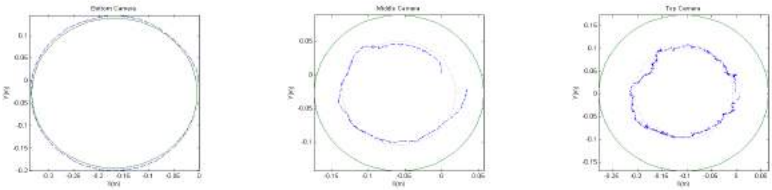


Fig. 10. Calculated ground truth (—), tracking trajectory (- - -) and circle fitting (···) result for three cameras mounted at different heights (0.44 m, 0.75 m, 1.33 m) from wooden floor during a circular path drive of the robot.

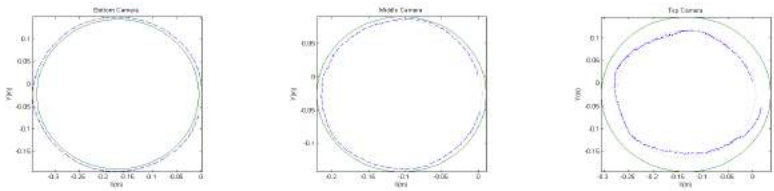


Fig. 11. Calculated round truth (—), tracking trajectory (- - -) and circle fitting (···) result for three cameras mounted at different heights (0.44 m, 0.75 m, 1.33 m) from mosaic floor during a circular path drive of the robot.

advance over present systems to use the existing environment itself as calibration pattern. Another advantage is its speed. The full calibration can be done under three minutes, unlike manual method which are much more slower. Furthermore, the robot can check its calibration at any time. Future works will be the studding effect of the intrinsic calibration of the 3D camera on the method and refinement of the method to live camera calibration during SLAM without need for any predefined paths.

References

[1] F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard, "3-d mapping with an rgb-d camera," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 177–187, 2014.

[2] T. Fulhammer, R. Ambru, C. Burbridge, M. Zillich, J. Folkesson, N. Hawes, P. Jensfelt, and M. Vincze, "Autonomous learning of object models on a mobile robot," *IEEE Robotics and Automation Letters*, vol. 2, no. 1, pp. 26–33, 2017.

[3] K. Tateno, F. Tombari, and N. Navab, "Real-time and scalable incremental segmentation on dense slam," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 4465–4472.

[4] M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger, "Real-time 3d reconstruction at scale using voxel hashing," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 169:1–169:11, Nov. 2013.

[5] C. C. Wang, "Extrinsic calibration of a vision sensor mounted on a robot," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 2, pp. 161–175, 1992.

[6] M. Pollefeys, R. Koch, and L. V. Gool, "Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 1998, pp. 90–95.

[7] S. J. Maybank and O. D. Faugeras, "A theory of self-calibration of a moving camera," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 123–151, Aug 1992.

[8] Q.-T. Luong and O. Faugeras, "Self-calibration of a moving camera from point correspondences and fundamental matrices," *International Journal of Computer Vision*, vol. 22, no. 3, pp. 261–289, Mar 1997.

[9] S. N. Sinha, M. Pollefeys, and L. McMillan, "Camera network calibration from dynamic silhouettes," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, 2004, pp. I–195–I–202 Vol.1.

[10] G. Carrera, A. Angeli, and A. J. Davison, "Slam-based automatic extrinsic calibration of a multi-camera rig," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 2652–2659.

[11] S. Miller, A. Teichman, and S. Thrun, "Unsupervised extrinsic calibration of depth sensors in dynamic scenes," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 2695–2702.

[12] S. Li, P. N. Pathirana, and T. Caelli, "Multi-kinect skeleton fusion for physical rehabilitation monitoring," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 5060–5063.

[13] J. Levinson and S. Thrun, *Unsupervised Calibration for Multi-beam Lasers*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 179–193.

[14] W. Maddern, A. Harrison, and P. Newman, "Lost in translation (and rotation): Rapid extrinsic calibration for 2d and 3d lidars," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 3096–3102.

[15] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, "Automatic targetless extrinsic calibration of a 3d lidar and camera by maximizing mutual information," in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, ser. AAAI'12. AAAI Press, 2012, pp. 2053–2059.

[16] J. Prankl, A. Aldoma, A. Svejda, and M. Vincze, "Rgb-d object modelling for object recognition and tracking," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 96–103.

[17] C. Tomasi and T. Kanade, "Detection and tracking of point features," International Journal of Computer Vision, Tech. Rep., 1991.

[18] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[19] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[20] D. Eberly, "Least squares fitting of data," *Chapel Hill, NC: Magic Software*, 2000.

[21] (2017) Point cloud library. [Online]. Available: http://pointclouds.org/about

[22] (2015) The vision4robotics library. [Online]. Available: https://www.acin.tuwien.ac.at/en/vision-for-robotics/software-tools/v4r-library/

[23] H. Haggag, M. Hossny, D. Filippidis, D. Creighton, S. Nahavandi, and V. Puri, "Measuring depth accuracy in rgbd cameras," in *2013, 7th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2013, pp. 1–7.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Towards a Definition of Educational Robotics

Julian M. Angel-Fernandez[1] and Markus Vincze[1]

***Abstract*— There is an increasing number of articles, web pages, robotic kits and other materials that are using the term Educational Robotics (ER) to refer to the use of robots in education, however the current definition of ER is still vague and open to misinterpretation. Therefore, anyone can claim that their work falls in the category of ER just because robots are involved. Despite all benefits of robotics, its incorrect use may be counterproductive. Therefore, the incremental use of the term ER is meaningless if it is not used correctly. Consequently, a concrete and precise definition of ER is required to support the development of it. This paper presents a first attempt to develop a concrete definition of ER, which describes all fields of study that constitutes it and how they are related between them. The definition is the result of the experience acquire during the participation of the European project Educational Robotics for STEM (ER4STEM).**

I. INTRODUCTION

Robotics has been mentioned by many researchers as a technology with significant potential to impact education [1], [2], [3], [4]. This is reflected in the increasing number of articles that uses the words *robotics* and *education* together, such as is presented in Figure 1-a. Likewise, the use of Educational Robotics (ER) has increased in the last two decades, as it is presented in Figure 1-b. Despite its increment, there is not a clear definition of what ER is and in many situations is mentioned just as a tool used in education [5], [6], [7] or as a vehicle to think about teaching, learning and education at large [8]. If ER is a merely tool, then several questions arise: What is robots' role in this "*tool*"? Who is responsible to develop further this "*tool*"? Is there any difference between educational robotics, educational robots, robots in education and robots for education? On the other hand, if it is seen as a vehicle: who has created the vehicle? How should the vehicle look like? How is it used?

While these and other questions are still open, it is difficult to correctly coordinate and establish criteria to identify works that can be categorized as ER. For example in the work 5-Step Plan [9], the researchers categorized their work as ER. However, they suggest that students are product designers that have to conceptualize a robot from scratch, without time nor knowledge to implement it. Then, participants could let their imagination go wild and come with creativity designs and tasks for their robots. In this case, robotics is used as a word to attract people's interest and not as a device to improve the learning experience. As a consequence, it could not be considered as ER because no robot is used to explain new concepts or strength others. Instead, it could be classified

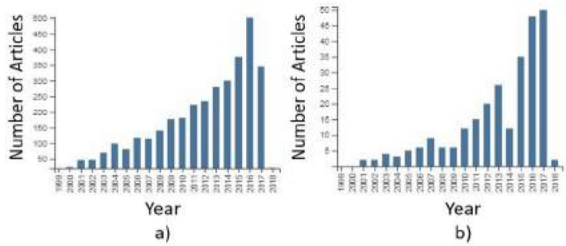


Fig. 1. Total number of articles per year retrieved from *Web of Science* using the following queries: a) "*Education* AND Robot**", which retrieves articles that contains any word derived from education and robot words, such as educational, robots, robotics, just to mention a few. Other words could be between these words and the order in which they appear does not matter. b) "*Educational Robotics*", which retrieve articles that contains the exact match of the words without other words in the middle.

as product design activity because participants learned the steps to correctly conceived and design a product, in this case a robot. However, this type of activities can create false or unreachable expectations of robots, which could frustrate people because current robots could not fulfill them. This frustration negatively affects the level and quality of the effort that people put into learning [10].

Despite all the benefits that robotics could have in fostering digital skills (e.g. programming [11]), STEM (e.g. Physics [12] and Mathematics [13]) and soft-skills (e.g. Creativity [14]), its incorrect use may be counterproductive [15] and it could stop its implementation in formal education settings (e.g. Schools). Therefore, a concrete definition, that specifies the meaning of ER is mandatory to correctly make move towards the right direction. This paper presents a first attempt to develop a concrete definition of ER, which describes components that constitute it and how they are related within them. The presented components are the result of the experience acquire during the participation in the European project Educational Robotics for Science, Technology, Engineering and Mathematics (ER4STEM) [1], which aims at realizing a creative and critical use of ER to maintain children's curiosity in the world.

This paper is organized as follows. Section II presents some of the related in ER and Robotics in Education. With this as a base line and in order to have a better understanding of ER, an analysis of stakeholders involved in ER is presented in Section III. Considering their requirements and needs, Section IV describes the ideal activity

[1]Julian M. Angel-Fernandez and Markus Vincze are with the Automation and Control Institute, Vienna University of Technology, Vienna, Austria {Angel-Fernandez, Vincze}@acin.tuwien.ac.at

[1]www.er4stem.com

in ER. Section V introduces the framework developed in ER4STEM, which aims to guide stakeholders in the design, implementation and evaluation of activities in ER. Based on the information presented, Section VI shows the field of studies that converge in ER and the definition of ER is provided. Finally conclusions are presented in Section VII.

II. Related Work

Robotics is used in different settings and platforms. Sullivan and Bers [16] studied how robotics and computer programming could be used from pre-kindergarten to second grade classrooms and what children could learn from them. They developed an eight week curriculum focused on teaching foundations of robotics and programming concepts. The robotic platform KIWI was used, which was specifically designed for young children (four years and up). The main particularities of KIWI are that it could be programmed using the Creative Hybrid environment for computer Programming (CHERP) and it does not require any computer to be programmed [17]. Similarly, Stoeckelmayr et al. [18] created eight workshops to introduce robotic concepts to kindergarten students using BeeBot. These workshop are created from their experience in Robocup Junior.

Robotics can be also used to teach physics and mathematics. For example Church et al. [19] created and implemented activities to explain acceleration, speed, harmonic motion, pendulums and sound's variables. Ashdown and Doria [12] used robots to introduce the Doppler effect. Their results suggest that participants engaged with the activity and they learned about the proposed topic.

In the last decade, researchers have come with the idea of using social robots in schools. Some researchers have investigated the features that a robot should have when is placed in a classroom [20]. They identified that motion is important for the participants, because it helps to break the monotony of classroom. Moreover participants highlight the importance of visualizing geometrical concepts in the real world and their interest in interacting with the robot in pet-like way. Other researchers focused on the impact of verbal cues given by a robot to participants [21], suggesting that it has a positive impact. Likewise Castellano et al. [22] as shown that people prefer robots that show empathy. These works are led by the Human Robot Interaction (HRI) community with especial focus on the social aspects of autonomous robots to improve the experience instead of the correct use of robots in education.

Despite the versatility of robotics in terms of topics, ages and situations, there is a missing understanding of ER to draw guidelines, scope and objectives. Without this understanding the real potential of robotics in education will not be completely unleashed and in some occasions, it could jeopardize the learning experience [15].

III. Stakeholders in Educational Robotics

In order to identify components that constitute ER to create a definition, it is required to understand who are the people involved, stakeholders, on it. ER4STEM's researchers [23] identified as stakeholders in ER:

- *Young people* are the ones who participate in ER activities offered by schools or other organizations. They are directly impacted by ER because are the ones who will participate in the activities created in ER.
- *Young people parents* may encourage their offspring to participate in activities or may not. Some parents would not be aware of the importance of digital skills [24], then they do not have any motivation to expose their offspring to activities that could foster them. This is an additional difficulty on the implementation of any ER because some parents would be hesitant to invest money and time.
- *Schools* are the place where formal teaching occurs and inside them two different stakeholders are present. (1) Teachers have as main responsibility to teach through the use of different methodologies. Although they are aware of the importance of Information and Communication Technologies (ICT) skills for teaching and new technologies [25], they are not confident about their knowledge in technology and its correct use in the classroom. (2) School boards or senior management decides over budget and established standards. They are influenced by the policymakers, government and parents.
- *Organizations offering educational robotics* which would be non-profit organizations offering ER activities, organizations based on profit or mixed versions (e.g. Clubs, projects, initiatives, universities, science and technology institutes). The activities offered by these organizations reach a wide audience and can create a big impact. Usually the activities offered by these institutions are considered as non-formal because are not link with any school curricula.
- *Universities* study, envision and developed technologies and techniques to be used in different fields, such as education and robotics. There are several stakeholders inside them that contribute in ER: educational researchers, teacher educators, engineering scientists and people involved in outreach programs. In many cases there is not much communication between them, which hinder the potential of ER.
- *Industry* is directly affected by people's skill sets and education. The demand in high quality knowledge workers in STEM fields is increasing worldwide but young people choosing STEM fields are not matching these numbers in demand [26]. There are even initiatives started by industry to counter these developments.
- *Educational Policy makers* are governmental organizations established with the purpose to lead the future of education.

ER does not have to address all of these stakeholder at once because covering all their requirements is a difficult task. Instead, ER, as a first step, must focus on those who have a direct impact on the quality of the activities, which

results could be used to support the investment on robotics. Consequently teachers, researchers, organizers of educational activities and industry have been identified as direct stakeholders [27]. They have different requirements from ER based on their needs and activities done by them [27], which are presented in Table I. All stakeholders do workshops. Teachers, researchers and organizers do activities, where they present information. Just teachers and researchers do research, and just teachers do lessons in schools. Regarding stakeholders' requirements of the activities, it is shown that most of the cases they require a good description of the activity to implement it. Just teachers and researches need activities that could be compared. On the other hand just teachers and organizers required activities that could be sustainable for long periods. The case of industry is particular because they required activities that let them promote their technologies. Although these stakeholder share some activities and their needs could be consider as complementary, there is not a good collaboration between them.

To exemplify the lack of collaboration, let's consider the case of researchers from all fields that ER converge to. In the ideal case, researchers communicate and establish common goals that are achieve through continue interaction within them. This produces ideas for new technologies and pedagogical approaches that could be used in education, which is reflected in the creation of workshops and lessons. These activities are expected to be described in enough detail that other people outside the group of work could implement. This provides several benefits: validate results, extend research beyond the original environment and use on different settings. Once the activity has been completed, researchers analyze the information collected, which brings new questions and suggestions for pedagogy and technology. Using these results as a base, researchers begin again with the cycle. However, the reality is that this collaboration between researchers is still limited or inexistent. Lets consider robotics and education researchers. In the ideal situation, they would work together to complement each other. Robotics researcher will provide the technological expertise that educational researches do not have and educational researchers will provide the knowledge to include the educational component during the design and development of robots and technologies. However, in many cases, this collaboration has been limited or inexistent.

IV. WHAT ARE THE ACTIVITIES COVERED IN EDUCATIONAL ROBOTICS?

The study done by ER4STEM's researches on the available literature found several weaknesses on how works on ER are documented [23]. (1) There is not a clear evidence how pedagogical theories were considered during the design of the activity. (2) Activities reported in many cases are not fully described, which limit their replicability. The last situation even occurs among researchers, who do not provide a detail description of their settings such as the ones reported in [16], [17], [18], [19]. In most of the cases researchers implemented as a workshops, which usually are done as extracurricular or non-formal activity. Therefore, researchers do not include learning outcomes and evidence of learning. In other cases they are implicit but not correctly documented. As a consequence, ER4STEM's researchers suggested that workshops and lessons must be treated as similar because regarding the place where the activity is implemented is required to have a clear learning outcomes and evidence of learning. This has several benefits. (1) The activities designed and implemented as a workshop are easily implemented as a lessons. This is due the description of objectives and proof of learning, which makes easer to recognize the connection with the school's curricula. (2) The evidence of learning let people to verify if the activity achieved the expected results or not. Also it could be used to measure the real impact of ER in the short term, which is important because it has not been quantified yet [28] and it would generate arguments towards the implementation of ER activities.

Based on all of these, ER4STEM's researchers suggest to call activities done in ER as pedagogical activities with the following characteristics: (1) clear learning outcomes and evidence of learning, which could be formal (e.g. assessment) or informal (e.g. write to a friend about what you have done today). (2) Use of one or more pedagogic methodology during the activity, which has to be described for each action in the activity. This is really important because technology alone is not enough to obtain desire learning outcomes [29]. (3) Description of the activity using an activity template (e.g. ER4STEM's activity template [30]). This will help other stakeholders to have a clear idea of all considerations taken into account and the assumptions done by the designer.

V. ER4STEM FRAMEWORK

General speaking, stakeholders are on their own when they have to design and implement a pedagogical activity in ER. Therefore, a person must have high knowledge in technology and education to correctly implement them. However, few people have all of this knowledge. As a consequence, ER4STEM is developing a framework that will guide any stakeholder on the design or adaptation, implementation and evaluation of pedagogical activities in ER. This is achieved through the explicit connection among pedagogical methodologies, knowledge in robotics and other areas, and 21st century skills [31].

The ER4STEM's framework provides four components. (1) An ontology of ER, which provides specific definition of words used in the field and connection between them. (2) Activity blocks, which are piece of activities that have been proven to be useful to foster specific skills and could be connected with other blocks to create a pedagogical activity. (3) Best practices, which are described from a literature reviewed done for creativity, collaboration, communication, critical thinking, evidence of learning, mixed gender teams, multiple entry points, changing and sustaining attitudes to STEM, and differentiation. (4) Processes for workshops and conferences for young people, which are based on the macro-process depicted in Figure 2.

TABLE I

ACTIVITIES AND NEEDS FOR EACH STAKEHOLDER WHO HAS A DIRECT IMPACT ON THE QUALITY OF ER'S ACTIVITIES [27].

	Teachers	**Researchers**	**Organizers of Educational Activities**	**Industry**
Activities	• Workshop • Presentation • Research • Lesson	• Workshop • Presentation • Research	• Workshop • Presentation	• Workshop
Requirements	• Pedagogical informed description • Compare activities and results • Well described activities • Sustainable activities	• Pedagogical informed description • Compare activities and results • Well described activities	• Well described activities • Sustainable activities	• Specific set of skills • Promote their technologies

The macro-process is compound by four main macro phases. (1) The first macro phase is divided in two possible steps, which represents the possibility to design an activity from scratch or adapt one from other existing activities. (2) Implementation macro-phase focuses on considerations involving the settings and the context in which the activity is going to take place. (3) Evaluation macro-phase focus on evaluating the implementation. (4) Improvement macro-phase focuses on possible improvements of the activity plan based on information derived from the implementation in real settings, on reflections from the teachers, the students and the designers. Once the activity has been improved, the cycle should be continuing with adapting the activity for future groups.

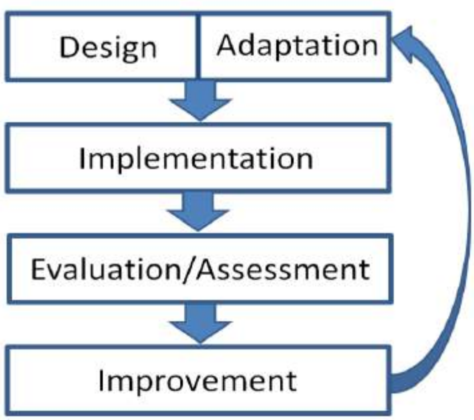


Fig. 2. Framework macro-process.

VI. WHAT FIELDS ARE INVOLVED IN EDUCATIONAL ROBOTICS?

Based on the information provided until this point, it is possible to observed certain fields of study that are involved in ER. Figure 3 presents a simplified view of them and their interconnections. By simplified, it is meant that just general fields are depicted and other fields (e.g. artificial intelligence) are omitted, without undervalue their contribution, to increase the clarity. Three main fields are presented in the figure. (1) Education embraces all sub-fields that are related to the study and improvement of learning experiences of people at all levels, from early childhood to university. (2) Robotics is the field that studies and improve robots. A tangible result is robotics platforms that in some cases have been used in education. A good example is the robotics platform Pioneer, which is meant to be used in research but it is also used in robotics college courses. This is called robotics in education. These platforms have been designed and implemented without considering their use in education. Therefore, they provide a hundred of functionality but there is not much space to create basic activities with them, which is called in education as black box [32]. (3) Human Computer Interaction (HCI) is a field that studies the interaction between computer and humans, aiming to improve user experience. This field has shown the importance of considering humans in the design of robotics platforms. As a result the field of Human Robot Interaction (HRI) was established, which is dedicated to understand, design and evaluate robotic platforms to be used with or by humans [33].

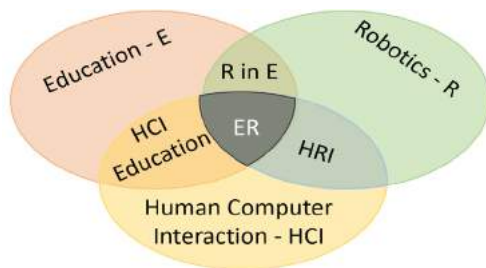


Fig. 3. A simplified view of fields of study that conformed educational robotics. Educational Robotics is the intersection between Education, Robotics and Human Robot Interaction. E means Education, R robotics, HCI Human Computer Interaction, HRI Human Robot Interaction, R in E Robots in Education, and ER Educational Robotics.

With all of this information and analysis, it is possible to conclude that ER is not just a tool but rather a field of study by its own, where many fields of study converge. Therefore, the definition of ER proposed in this article is the following:

Educational Robotics is a field of study that aims to improve learning experience of people through the creation, implementation, improvement and validation of pedagogical activities, tools (e.g. guidelines and templates) and technologies, where robots play an active role and pedagogical methods inform each decision.

It is important to highlight that this definition covers existing categories of the use of robotics in education. Alimisis and Kynigos [4] identified two categories. (1) robotics as learning object focuses on robotic related topics, such as computer vision and artificial intelligence. (2) Robotics as learning tool sees robots as a tool to teach other subjects, such as science or math. Eguchi has proposed a third category [5] that sees robots as learning aids, which would be in most of the cases social robots, such as the Robot-Tutor in collaborative learning scenarios [34] and Robot-Tutor in teaching languages [35]. Robotic platforms in the first two categories are characterized to be cheap, and with limited number of sensors, actuators and computer processing, in comparison to its industrial counterparts. Also they are not limited to traditional programming languages (e.g. Python, C++ and C) but they used novel programming languages to improve the learning experience (e.g. Scratch [36] and tangible programming [37]). Robots in the last category are expensive due to they have to interact in a natural way with humans and behave in a way that is comfortable for humans.

VII. CONCLUSIONS

This paper presented stakeholders and the requirements of teachers, researchers, workshops organizers and industry in ER. These requirements were used to draw the components of an activity in ER. It was suggested that there should not be difference between lessons and workshops because both must have learning outcomes and proof of learning. These would enable stakeholder to use these activities designed and implemented for workshops in lessons and vice-versa. Also, it would allow to measure the impact of robotics in education, which is still unknown [28]. Therefore the use of the tag pedagogical activity was suggested to name activities in ER, which have the following characteristics: clear learning outcomes and evidence of learning, use of one or more pedagogic methodology, and description of the activity using specific templates template(e.g. ER4STEM's activity template [30]). Also, it was presented the ER4STEM framework, which aims to guide any stakeholder on the design or adaptation, implementation and evaluation of pedagogical activities in ER. Based on all these information, it was presented the fields that converge in ER and it was suggested the following definition for ER:

Educational Robotics is a field of study that aims to improve learning experience of people through the creation, implementation, improvement and validation of pedagogical activities, tools (e.g. guidelines and templates) and technologies, where robots play an active role and pedagogical methods inform each decision.

This definition covers existing categories of the use of robotics in education: robotics as learning object [4], robotics as learning tool [4], and as leaning aid [5].

The authors hope that these definitions are used as a based to define the field of ER and the characteristics of the activities developed on it. These clear definitions will help different stakeholders to understand and apply correctly the knowledge created in the field and to strength the collaboration between different researchers and even stakeholders.

ACKNOWLEDGMENT

This work is funded by the European Commission through the Horizon 2020 Programme (H2020, Grant agreement no: 665972). Project Educational Robotics for STEM: ER4STEM. The authors would like to to thank Georg Halmetschlager for his comments.

REFERENCES

[1] S. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*. New York, NY, USA: Basic Books, Inc., 1980.
[2] D. Catlin and M. Blamires, "The principles of educational robotic applications (era)," in *Constructionism Conference, Paris*, 2010.
[3] M. G. da Silva Filgueira and C. S. González González, "Pequebot: Propuesta de un sistema ludificado de robótica educativa para la educación infantil," 2017.
[4] D. Alimisis and C. Kynigos, "Constructionism and robotics in education," *Teacher Education on Robotic-Enhanced Constructivist Pedagogical Methods*, pp. 11–26, 2009.
[5] A. Eguchi, "Educational robotics theories and practice: Tips for how to do it right," *Robotics: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, vol. 193, 2013.
[6] S. Atmatzidou and S. Demetriadis, "A didactical model for educational robotics activities: A study on improving skills through strong or minimal guidance," in *International Conference EduRobotics 2016*. Springer, 2016, pp. 58–72.
[7] F. Agatolio, M. Pivetti, S. Di Battista, E. Menegatti, and M. Moro, "A training course in educational robotics for learning support teachers," in *International Conference EduRobotics 2016*. Springer, 2016, pp. 43–57.
[8] D. Alimisis, "Robotics in education & education in robotics: Shifting focus from technology to pedagogy," in *Proceedings of the 3rd International Conference on Robotics in Education*, 2012, pp. 7–14.
[9] L. Lammer, A. Weiss, and M. Vincze, "The 5-step plan: A holistic approach to investigate children's ideas on future robotic products," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. ACM, 2015, pp. 69–70.
[10] C. Wegner, F. Strehlke, and P. Weber, "Investigating the differences between girls and boys regarding the factors of frustration, boredom and insecurity they experience during science lessons," *Themes in science and technology education*, vol. 7, no. 1, pp. 35–45, 2014.
[11] D. Tochek, J. Lape, and V. Fuglk, "Developing technological knowledge and programming skills of secondary schools students through the educational robotics projects," *Procedia - Social and Behavioral Sciences*, vol. 217, pp. 377 – 381, 2016.
[12] J. Ashdown and D. Doria, "A robotics based design activity to teach the doppler effect." in *IEEE 2nd Integrated STEM Education Conference*. IEEE, 2012.
[13] S. Hussain, J. Lindh, and G. Shukur, "The effect of lego training on pupils' school performance in mathematics, problem solving ability and attitude: Swedish data," *Journal of Educational Technology & Society*, vol. 9, no. 3, 2006.
[14] J. M. Angel-Fernandez and M. Vincze, "Introducing storytelling to educational robotic activities," in *Proceedings of EDUCON2018 - IEEE Global Engineering Education Conference*. IEEE, 2018.
[15] J. Sharkey, "Establishing twenty-first-century information fluency," *Reference & User Services Quarterly*, vol. 53, no. 1, p. 33, 2013.
[16] A. Sullivan and M. U. Bers, "Robotics in the early childhood classroom: learning outcomes from an 8-week robotics curriculum in pre-kindergarten through second grade," *International Journal of Technology and Design Education*, vol. 26, no. 1, pp. 3–20, 2016.

[17] A. Sullivan, E. R. Kazakoff, and M. U. Bers, “The wheels on the bot go round and round: Robotics curriculum in pre-kindergarten,” *Journal of Information Technology Education*, vol. 12, 2013.

[18] K. Stoeckelmayr, M. Tesar, and A. Hofmann, “Kindergarten children programming robots: a first attempt,” *Proc. Robotics in Education*, pp. 185–192, 2011.

[19] W. J. Church, T. Ford, N. Perova, and C. Rogers, “Physics with robotics - using lego mindstorms in high school education.” in *AAAI Spring Symposium: Educational Robotics and Beyond*. AAAI, 2010.

[20] E. Walker and W. Burleson, “User-centered design of a teachable robot,” in *International Conference on Intelligent Tutoring Systems*. Springer, 2012, pp. 243–249.

[21] L. N. Brown and A. M. Howard, “The positive effects of verbal encouragement in mathematics education using a social robot,” in *Integrated STEM Education Conference (ISEC), 2014 IEEE*. IEEE, 2014, pp. 1–5.

[22] G. Castellano, I. Leite, A. Paiva, and P. W. McOwan, “Affective teaching: learning more effectively from empathic robots,” 2012.

[23] J. M. Angel-Fernandez, L. Lammer, C. Kynigos, I. Gueorguiev, P. Varbanov, W. Lepuschitz, A. Duca, J. Pullicino, M. Grizioti, S. Nikitopoulou, C. Girvan, and P. Vrba, “Best practice and requirements,” TU Wien, University of Athens, ESI, Cardiff University, AcrossLimits and Certicos,” Deliverable, 2016.

[24] F. E. Digital, “Competencias digitales en españa como mejorarlas?” Internet, www.espanadigital.org, 2015.

[25] Eurydice, “The teaching profession in europe: Practices, perceptions, and polices,” Internet, 2015.

[26] M. Caprile, R. Palmén, P. Sanz, and G. Dente, “Encouraging stem studies: Labour market situation and comparison of practices targeted at young people in different member states,” *European Parliament’s Committee on Employment and Social Affairs*, 2015.

[27] J. M. Angel-Fernandez, C. Kynigos, W. Lepuschitz, J. Pullicino, M. Grizioti, C. Girvan, and C. Todorova, “Towards an extended definition of er4stem framework,” TU Wien, University of Athens, ESI, Cardiff University and AcrossLimits,” Deliverable, 2017.

[28] F. B. V. Benitti, “Exploring the educational potential of robotics in schools,” *Comput. Educ.*, vol. 58, no. 3, pp. 978–988, Apr. 2012.

[29] C. Richards, “Towards an integrated framework for designing effective ict-supported learning environments: the challenge to better link technology and pedagogy,” *Technology, Pedagogy and Education*, vol. 15, no. 2, pp. 239–255, 2006.

[30] N. Yiannoutsou, S. Nikitopoulou, C. Kynigos, I. Gueorguiev, and J. A. Fernandez, “Activity plan template: a mediating tool for supporting learning design with robotics,” in *Robotics in Education*. Springer, 2017, pp. 3–13.

[31] P. P. for 21st Century Learning, “Framework for 21st century learning,” Internet, http://www.p21.org/our-work/p21-framework.

[32] C. Kynigos, “Black-and-white-box perspectives to distributed control and constructionism in learning with robotics,” in *Proceedings of SIMPAR workshops*, 2008, pp. 1–9.

[33] M. A. Goodrich and A. C. Schultz, “Human-robot interaction: A survey,” *Found. Trends Hum.-Comput. Interact.*, vol. 1, no. 3, pp. 203–275, Jan. 2007. [Online]. Available: http://dx.doi.org/10.1561/1100000005

[34] P. Alves-Oliveira, S. Janarthanam, A. Candeias, A. Deshmukh, T. Ribeiro, H. Hastie, A. Paiva, and R. Aylett, “Towards dialogue dimensions for a robotic tutor in collaborative learning scenarios,” in *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*. IEEE, 2014, pp. 862–867.

[35] T. Belpaeme, P. Vogt, R. van den Berghe, K. Bergmann, T. Göksun, M. de Haas, J. Kanero, J. Kennedy, A. C. Küntay, O. Oudgenoeg-Paz, *et al.*, “Guidelines for designing social robots as second language tutors,” *International Journal of Social Robotics*, 2017.

[36] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, and Y. Kafai, “Scratch: Programming for all,” *Commun. ACM*, vol. 52, no. 11, pp. 60–67, Nov. 2009.

[37] M. U. Bers and M. S. Horn, “Tangible programming in early childhood,” *High-tech tots: Childhood in a digital world*, vol. 49, pp. 49–70, 2010.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 innsbruck university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Safety of Industrial Applications with Sensitive Mobile Manipulators – Hazards and Related Safety Measures

Andreas Schlotzhauer[1], Lukas Kaiser[1] and Mathias Brandstötter[1]

Abstract— **The areas of application of robot systems are gradually expanding and mobile manipulation is an important and consistent further development for industrial applications. Although human-robot interaction with these systems becomes easier, the mechatronic design, the integration and safety regarding real applications remain challenging. This paper describes identified dangers and possible hazards of industrial mobile robot systems and sensitive mobile manipulators. Based on a study of advanced sensor technologies and safety concepts, solutions and measures for risk reduction are proposed to counteract these risks. As a key element in mobile robotics, common drive architectures are evaluated with regard to their impact on the general application safety.**

I. INTRODUCTION

With the focus on Industry 4.0 and the associated increasing digitization of the supply chain, there is a high demand for versatile tools in the manufacturing industry [1], [2]. This development includes robot systems that can be used flexibly in such environments. The relatively new field of sensitive mobile manipulation has evolved through major advances in technology and the related development of collaborative manipulators, which fills an aspect of these needs. Such robotic systems have to satisfy a multitude of basic requirements and general conditions, which are examined in this work.

A. Abilities of Mobile Manipulators

Mobile manipulators, sometimes simply called mobile robots, are the fusion of sensitive manipulators and mobile platforms. Therefore, they combine the two major advantages of both technologies: (i) the capability of working in close proximity to the human, which enables collaboration, and (ii) autonomous relocation and adaptation to a changing environment, which results in novel industrial applications like discussed in [3] and [4].

B. Norms and Standards

The safety requirements for all types of machinery, within the European Union, is regulated by the so called *Machinery Directive* [5]. The ISO 12100 [6] is harmonised with the directive and gives general guidance for the safety throughout the life cycle of a machine. One of the main aspects of these documents is the risk-assessment and -reduction of a machine before first operation. The ISO 10218 [7], [8] extends the previously mentioned general standard and also includes more specific safety requirements tailored to the demands of industrial robot applications.

[1]All authors are with JOANNEUM RESEARCH ROBOTICS - Institute for Robotics and Mechatronics – Mechatronic Systems Group, Austria `<firstname.lastname>@joanneum.at`



Fig. 1. Mobile manipulator CHIMERA from JOANNEUM RESEARCH

The close vicinity between the robot and the operator in collaborative applications yields new and higher risks. For this reason, the International Organization for Standardization (ISO) released a Technical Specification ISO/TS 15066 [9] addressing the special issues of collaborative robots. For *Automated Guided Vehicles* (AGVs) on their own two rather old European standards ([10], [11]) exist. These standards are currently revised by the ISO to form the ISO/DIS 3691-4 [12]. Moreover, at the moment there is only one active standard [13] that directly considers the overall system of a mobile manipulator, but only in the context of personal care and therefore excluding industrial use.

To fill this gap, the sub-committee R15.08, of the American *Robotic Industries Association* (RIA), is currently developing a new standard for *"Mobile Robot Safety"*. At this time, the developers and integrators of such system are responsible to go beyond the current standards to make their products as safe as possible.

C. Market Overview

As mentioned above, the mobility of a mobile manipulator is one of its key advantages over conventional mobile industrial robots. To be able to operate on the shop floor, next to and hand in hand with human workers, the requirements for localization and navigation are high. Although all mobile robots on the market have some kind of these features, it is the quality that sets them apart. Other distinguishing features are, for instance, the maximum loading capacity, runtime, charging time, travel speed, and the quality of

maps created by integrated SLAM algorithms. In terms of the safety relevant wheel configuration, there are no major differentiations, as most systems use either a differential drive with additional castors or four omnidirectional wheels for increased stability (see section III-A). Examples for the various platform types are, e.g., CHIMERA by JOANNEUM RESEARCH (see Fig. 1) that is based on a differential drive and the *KMR iiwa* by KUKA as an example for a platform with an omnidirectional drive.

II. HAZARDS RELATED TO MOBILE MANIPULATION

There are some common hazards that can occur in every electro-mechanical system, like sharp edges, collision by moving machinery parts, the chance of getting in contact with high voltage or hot/cold surfaces, as well as loud noise, radiation or vibration (for more details see [6]). In the following special hazardous situations are discussed, that can occur only or especially in industrial robotic applications with mobile manipulators.

A. Hazardous Situations

One source of danger is the movement of a mobile manipulator, more specific the movement of the mobile platform, the attached manipulator or both together. A major risk is the collision (transient and quasi-static contact) with a human, which can only be managed with supplementary measurements.

Another challenge and source of danger is the stability of the whole robot during driving and handling of objects. Especially when the mobile platform and the manipulator are moving simultaneously, the dynamic of the whole robot needs to be taken into account. When the movement of the robot is limited (e.g., the robot fell over, the remaining stored energy is not sufficient, the drive is damaged), the robot should still be manually movable, as it could block an emergency escape route or be a barrier for other vehicles or humans. On the other hand the robot should not move unintentionally while being on an uneven surface or doing a precise task with its end-effector, as this could also lead to further accidents.

Even when the risk of a collision is reduced with sensors, there might still be a chance for hazardous situations. This could happen, e.g., when the robot converges to a docking station, its view is blocked by obstacles and objects are not visible from the robots point of view. This is also relevant for objects carried by the robot.

Dynamic changes in the environment and unknown objects in that environment could lead to situations, that where not predicted during integration and therefore, are not covered in the previously performed risk assessment. The interaction with dangerous objects/tools and the presence in unsuitable areas can hardly be completely excluded.

A communication between the robot system and humans tailored to the application should be considered to avoid confusion and misunderstanding and therefore to decrease the probability of the occurrence of a hazardous situation. More specifically, this means that one cannot assume that only qualified persons will interact with the robot (e.g., visitor groups that are guided through a production hall).

B. Possible Injuries

There is a wide scope of injuries that could occur due to the described hazardous situations. Special attention should be given to the high possibility of collisions between a robot and a human, as this is a unique property to collaborative robot applications. [14] studies possible soft tissue injuries and in [9] thresholds for the human experience of pain are given. For possible contact situations of an application, the compliance with these thresholds can be verified, to ensure the prevention of any injury [15].

III. DESIGN CONCEPTS

In order to design a robot system for industrial use, the safety aspects must be taken into account from the very beginning. Measures to reduce the risks, identified by a risk assessment [6], are grouped and prioritized into

1) Inherently safe design,
2) Complementary measures and
3) Organisational measures.

In the following, major concepts are presented to design and safeguard an industrial mobile application in different aspects.

A. Drive of Mobile Robots

An important element of a mobile platform or a mobile manipulator and therefore for the whole system is the locomotion mechanism, typically a drive, of the robot. If we restrict ourselves to wheeled mobile robots, then there are 4 basic types of wheels: standard wheels, castor wheels, Swedish wheels and spherical wheels. Each having different sliding and rolling constraint and affecting the maneuverability and controllability of the drive differently [16]. It is desirable, that the mobile manipulator is static stable and does not tilt during driving. A hyperstatic wheel geometry could lead to loose control on uneven floor. In general, omnidirectional drives in contrast to differential drives allow to react more flexible to dynamic changes in the planned path and are more suitable for narrow workspaces but could lead to undesired movements without active control (e.g. on an uneven floor). In table I common drives, presented in [16], are evaluated concerning their implication to safety, when used in a mobile industrial application. Also particular realizations can have different properties, and therefore, the general tendency is described.

B. Design of the Robot

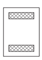
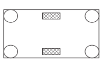
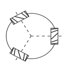
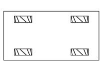
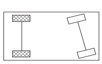
Concerning the mechanical design of the whole robot and end-effector, safety should be considered from the early beginning. The identified main mechanical design concepts related to safety are

- Lightweight design,
- Rounded edges,
- Compliant covering,

- Maximizing potential collision surfaces,
- Excluding bruise and shear of body parts,
- Limited workspace.

Also the physical interaction between the robot and a human is neither necessary nor scheduled or even prevented by supplementary measurements. It has to be noted, that a contact could still occur by intentional missus or due to a failure. In terms of safety (and efficiency) the whole robot should be as light as possible, especially all moving parts of the manipulator. For a good stability the center of gravity should be near the ground. Rounded edges and a soft covering can not only decrease the collision force and pressure, but also give a more comfortable feeling while touching the robot, which again could lead to higher acceptance by the operators. The bruise and shear of human limbs should be impossible anyway, not at least for the manual repair and maintenance of the robot. Rounded surfaces can also prevent that objects can be placed on the robot and fall down during driving. The area where the robot can move should not be larger than necessary and also the manipulator workspace can be limited if possible.

TABLE I
SAFETY OF COMMON DRIVES

	Wheel geometry	***Safety implications***
(a)		*Quasi-omnidirectional; Static instable, if the center of gravity is above the wheel axis;* Even the drive could be stabilized by a control, the inherent safety is low and the use is not recommended for industrial applications
(b)		*Quasi-omnidirectional; Hyperstatic; Static stable;* The differential drive enable precise path tracking. The risk of tilting is low but it is unsuitable for uneven floor. Standard wheels have a high payload and are robust.
(c)		*Omnidirectional; Static stable;* The ability to change the movement orthogonal to the moving direction can prevent hazardous situations. The chance of tilting is higher than in (b) and (d). The payload of Swedish wheels is in general lower than of standard wheels and the control is more vulnerable, which could be a problem for path tracking.
(d)		*Quasi-omnidirectional; Hyperstatic; Static stable;* The drive has similar safety implications than (c) but has in general a higher payload and the chance of tilting is less. When the swedish whells oriented the same, the platform can passively move in rolling direction of the small rollers.
(e)		*Not omnidirectional; Hyperstatic; Static stable;* This wheel geometry enable the most precise and robust path tracking, when using rigid axes and a Ackerman steering. The low maneuverability could lead to problems while moving the platform manually or navigating in small areas.

	Powered standard wheel
	Passive spherical wheel
	Passive standard wheel
	Powered Swedish wheel

C. Gripping

Especially the end-effector of a mobile manipulator, often a gripper, should be designed following the above listed principles, since in many cases the end-effector is the only physical interface with the environment (except the wheels) and the human. In mobile manipulation form-fit gripping should be preferred over force-fit gripping, as a gripped object cannot be lost after power loss, when it is slippery or even with a dynamic movement. By monitoring the gripping force and displacement of the gripper fingers, the compliance of the grasp object can be determined and the presence of human limbs can be detected. If gripping tasks or the handling of tools require fine mechanics or sharp edges, the covering or flexible suspension of the whole end-effector can be a solution [17]. In some applications flexible gripper fingers or a suctions cup can avoid sharp edged, but they might lack in precision and payload.

IV. AVAILABLE SAFETY TECHNOLOGY

As mentioned in the previous chapter, one possibility to reduce a risk is to put complementary measures into place. Historically, that is understood to putting the robot behind a ridged or light fence. With mobile manipulators this is usually not possible, and hence, the safety relies heavily on several modern sensor technologies which are presented in this section.

A. Localisation, Planning and Navigation

One key-aspect for safety in mobile robots is their capability of recognising their surroundings and acting to that accordingly. By constantly mapping its surrounding and localizing its position (*simultaneous localization and mapping* (SLAM)) the mobile robot is able to navigate safely in unstructured environments without collisions. To improve the localization and thereby also the safety, artificial features (bar codes, QR-tags, magnets etc.) can be used, although they might be covered by obstacles. On the other hand, natural landmarks are more challenging to detect, but with the advantage of lower risk of manipulation, damaging or covering.

B. Sensors

Regarding the risk, the sensor addresses, a suitable sensor technology has to be chosen. Different sensor types cover different aspects of the real world and can be distinguished by their robustness against external factors. In the following common sensor types and their impact on application-safety are presented.

1) Odometry: The major advantage of common odometry sensors like rotary encoders or accelerometers, is the high robustness, due to the basic underlying principles. Because the measurement is relative to the last information (except absolute rotary encoders), the error accumulate and the reliability of the sensor information decreases over time, which can be stabilized with extra reference points (global reference). When the position of the robot is derived from wheel rotation, slipping distorts the position accuracy until the next absolute reference. The computational power, needed for evaluation, is relatively low.

2) Tactile Sensing: A mechanical switch is very robust, although the derived information is simple. With higher complexity more sophisticated information can be captured. Besides the usage of tactile sensors in external input devices, artificial robot skins enable tactile sensing to standard robots, with the help of pressure sensitive air cushion or distributed and flexible force sensing elements on the robot surface. In that way the contact with humans can be detected and the avoidance of injuries is possible by appropriate reactions. The contact with the environment can also be perceived by force-torque measurements in the robot joints or base. This method might be difficult or even useless when a stationary robot is mounted on a mobile platform, without concerning and modeling the dynamic of the whole mobile manipulator.

3) Distance Sensing: There are several sensors available based on *Time Of Flight* (TOF) principle for measuring distances like SONAR, LIDAR and RADAR. Known issues with such systems are crosstalk, multi-reflections, absorption/permeability or insufficient reflection. Environmental conditions, e.g. sunlight or glass walls, can also decrease the performance.With a suitable arrangement of capacitive sensors, the orientation and distance to obstacles or people in the immediate vicinity can be calculated [18].

C. Sensor Fusion

The basic idea behind sensor fusion related to safety is to increase the coverage or integrity of the extracted information from different sensors by combining several sources of data. The combination of different types of sensors based on different operating principles decreases the chance of malfunction related to a common cause. This is also crucial for the redundancy requirements of safe interaction with the environment. In case of a mismatch between two channels the trustworthiness is not given any more for both signals and therefore the derived information as well. By cross checking more than two channels the failure of one specific signal can be recognized with high probability. The difference between channels can also be caused by the limited capabilities of different operating principles, e.g., detecting a pane of glass with an optical sensor versus an ultrasonic sensor. This reduces the trustworthiness of the consolidated sensor data but increases the scope of perceivable information. It is not trivial to distinguish between these two situations, however it can be achieved by pairing two similar sensors for each operating principle.

D. Safety of dynamic workflow

Due to undetermined dynamic changes in flexible mobile robotic applications, not every possible situation can be analysed regarding its risk beforehand. Therefore some kind of dynamic risk analysis during runtime would be beneficial. To realize this approach some kind of intelligent system is necessary to be aware of the situation and assess the same. Image classification/object recognition is widely used to achieve this goal. Neuronal networks are able to find dependencies within vast datasets (e.g., image collection) which can be used to evaluate new situations. This results in high level information that can not be derived from any other sensor with the drawback of not being replicable and therefore also not predictable, which is problematic for safety related functionalities. Potential fields with risk sources can be used to react and re-plan actions [19].

E. Multistage Safety Concept

As different types of sensors have different levels of reliability, a multistage concept can be used to increase productivity without sacrificing safety. Such systems could switch automatically between different safe modes depending on sensor input (e.g., distances) and the state of the available safety features (e.g., trustworthiness, failures), keeping the productivity as high as possible. For example, the use of an AI based vision system increases the predictability of the movement of humans, if the feature can not be trusted any more or fails, the system can then reduce the speed by relying on, e.g., the still working LIDAR scanners without beeing forced to stop the system.

V. CONCLUSIONS

The dissemination of flexible mobile application comes with chances and risks. While mobile manipulation is highly developed in research labs, industrial application remain tough, due to the lack of reference standards and experience-based knowledge. To face hazards in dynamic environments, a solid design that increases inherent safety is a fundamental requirement for a safe application. Also a suitable mechanical design is not enough. Instead, only advanced sensor technology or even AI-based methods can achieve a high level of safety, but in contrast they are error-prone and difficult to maintain. Redundancy and the combination of different technologies is crucial to overcome this problems. A good safety concept should not hinder the advanced possibilities of mobile manipulation, whereby operational safety should be in the foreground. To achieve this, knowledge of hazards and countermeasures must be transferred from the laboratory to the integrators and operators.

ACKNOWLEDGMENT

The results incorporated in this paper were gained within the scope of the project ”HRC-Safety for employees” commissioned by the Allgemeine Unfallversicherungsanstalt (AUVA).

REFERENCES

[1] H. Hirsch-Kreinsen, "Digitization of industrial work: development paths and prospects," *Journal for Labour Market Research*, vol. 49, pp. 1 – 14, 2016.

[2] D. Wurhofer, T. Meneweger, V. Fuchsberger, and M. Tscheligi, "Reflections on operators and maintenance engineers experiences of smart factories," in *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, 2018, pp. 284–296.

[3] K. Zhou, G. Ebenhofer, C. Eitzinger, U. Zimmermann, C. Walter, J. Saenz, L. P. Castao, M. A. F. Hernndez, and J. N. Oriol, "Mobile manipulator is coming to aerospace manufacturing industry," in *2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings*, Timisoara, Romania, Oct. 2014, pp. 94–99.

[4] FLEXIFF consortium. (2018, Mar.) Flexiff - flexible intralogistics for future factories. [Online]. Available: http://www.flexiff.at/

[5] "Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast)," The European Parliament and the Council of The European Union, Brussels, Belgium, 2006.

[6] "ISO 12100:2010-11 Safety of machinery – General principles for design – Risk assessment and risk reduction," International Organization for Standardization (ISO), Geneva, Switzerland, 2013.

[7] "ISO 10218-1:2011-07 Robots and robotic devices – Safety requirements for industrial robots – Part 1: Robots," International Organization for Standardization (ISO), Geneva, Switzerland, 2012.

[8] "ISO 10218-2:2011-07 Robots and robotic devices – Safety requirements for industrial robots – Part 2: Robot systems and integration," International Organization for Standardization (ISO), Geneva, Switzerland, 2012.

[9] "ISO/TS 15066:2016-02 Robots and robotic devices – Collaborative robots," International Organization for Standardization (ISO), Geneva, Switzerland, 2016.

[10] "EN 1525:1997-09 Safety of industrial trucks - Driverless trucks and their systems," European Committee for Standardization (CEN), Brussels, Belgium, 1997.

[11] "EN 1526:1997-09 Safety of industrial trucks - Additional requirements for automated functions on trucks," European Committee for Standardization (CEN), Brussels, Belgium, 1997.

[12] "ISO/DIS 3691-4 Industrial trucks – Safety requirements and verification – Part 4: Driverless industrial trucks and their systems," International Organization for Standardization (ISO), Geneva, Switzerland, 2018.

[13] "ISO 13482:2014-02 Robots and robotic devices – Safety requirements for personal care robots," International Organization for Standardization (ISO), Geneva, Switzerland, 2014.

[14] S. Haddadin, A. Albu-Schäffer, and G. Hirzinger, "Soft-tissue injury in robotics," in *Proceedings of the 2010 IEEE International Conference on Robotics and Automation*, 2010.

[15] "Kollaborierende Robotersysteme - Planung von Anlagen mit der Funktion Leistungs- und Kraftbegrenzung FB HM-080," Deutsche Gesetzliche Unfallversicherung (DGUV), Berlin, Germany, 2017.

[16] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to autonomous mobile robots*. MIT press, 2011.

[17] R. Weitschat, J. Vogel, S. Lantermann, and H. Höppner, "End-effector airbags to accelerate human-robot collaboration," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 2279–2284.

[18] M. Brandstötter, S. Mühlbacher-Karrer, D. Schett, and H. Zangl, "Virtual compliance control of a kinematically redundant serial manipulator with 9 dof," in *Advances in Robot Design and Intelligent Control*. RAAD 2016, 2016, pp. 38–46.

[19] B. Lacevic and P. Rocco, "Kinetostatic danger field - a novel safety assessment for human-robot interaction," in *International Conference on Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ*, Taipei, Taiwan, 2010, pp. 2169–2174.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
 ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

MMAssist_II: Assistance in production in the context of human machine cooperation

Christian Wögerer[1], Matthias Plasch[1], Manfred Tscheligi[2], Sebastian Egger-Lampl[2] and Andreas Pichler[1]

***Abstract*— MMAssist_II is a national Austrian flagship project for research, development and establishment of assistance systems which could be used as a tool box for different applications. Besides a fundamental understanding of demands for such assistance units also a demonstration in industrial near production settings including an extensive evaluation is part of the project. Therefore, a mighty consortium of 9 scientific partners and 16 Industrial partners was formed.**

I. INTRODUCTION

Initial situation: Austrian production companies manufacture goods of high quality and have a staff of well-trained employees. However, companies currently face technological and societal challenges to which they have to react to in order to continually provide competitive goods on an international level. These challenges include the demand of customers for individualized products, which leads to smaller lot sizes and faster production cycles. At the same time, production machines are more and more connected and equipped with sensors. This leads to an increased information density and more complexity for the workers, which induces a higher workload and stress. Furthermore, Austria is experiencing a demographic change. As Austrian citizens get older, they stay longer in employment. All of these trends, as well as the goal to keep up the high quality of produced goods, lead to an increased need of optimized assistance for the worker in the factory.

II. THE PROJECT MMASSIST_II

A. Key Facts

MMAssit_II was launched in May 2017 and will run until April 2020. The project involves 25 different partners from research and industry, which are key players for research and manufacturing in Austria [1], [2], [3]. The partners expertise covers the whole manufacturing value chain from basic research to industrial manufacturing of high tech products and services. This consortium was set up to have all necessary competences without any overlap in research, and besides technical capacities there is also social-economic knowledge available. The industrial partners cover a wide range of different technical branches and provide real use cases to demonstrate the results in a production near environment.

[1]Christian Wögerer, Matthias Plasch and Andreas Pichler are with ProFactor GmbH {firstname.lastname}@profactor.att

[2]Manfred Tscheligi and Sebastian Egger-Lampl are with the AIT {firstname.lastname}@ait.ac.at



Project- Facts

FFG Number.:	858623
Titel:	Assistenzsysteme in der Produktion im Kontext Mensch – Maschine Kooperation
Acronym:	MMAssist_II (www.mmassist.at)
Coordination:	PROFACTOR GmbH/ University Salzburg
Start:	01.05.2017
Duration:	36 months
Funding:	3 912 568 €
Total Costs:	6 131 405 €
Partner:	25 (10 scientific Partner, 15 Company partner)

Fig. 1. Key facts of the MMASSIST_II project

B. Objectives

The goal of the project partners in MMAssist_II is to explore assistance systems for employees in production environments and to develop these systems. This is necessary to overcome future technical and socio-technical challenges for production, by setting new paradigms of industrial assistance.



Challenges for future Production processes

Technical	Sozio Technical
Small lot sizes	Stress and work pressure
Common Setup and multiple machine operation	Physische Anstrengung
Physical effort	Knowledge Management
Flexible Assistant systems	High diversity of employees

Fig. 2. Challenges for Future production processes

OBJECTIVE 1: Exploration of modular, reusable assistance systems The project partners will develop assistance systems that can be used not only for the specific individual cases, but are applicable in different contexts and for different applications. The purpose is to establish a general approach for implementing assistance systems for employees in manufacturing companies.

OBJECTIVE 2: Context oriented detection of assistance needs Methods are developed, to enable the identification of the assistance needs of people in the vicinity of the machines from machine point of view. The purpose is to explore intelligent assistance systems, which offer targeted assistance only if it is needed.

OBJECTIVE 3: Improve the work and assistance expe-

rience As a major goal, the project partners will implement assistance systems that increase positive factors of work and assistance experience while they are used, and reduce negative factors. Thus, it will be achieved that the systems are accepted by users and contribute to an improvement of their daily work.

OBJECTIVE 4: Applicability in real production environments The project partner aims to use the implemented assistance systems application at the industrial partners production facilities and to evaluate in terms of productivity, acceptance through the staff and ergonomics. This evaluation should prove that the assistance systems are also usable in real production environments and beyond the project.



Fig. 3. Basic technologies avialable for MMAssist_II Assitance Units

III. BASIC TECHNOLOGIES

In the MMAssist_II project, nine scientific partners from Austria provide different basic technologies for various Assistance Units. Either these technologies are ready for implementation or they were developed ready for use. Most challenging problem are the interfaces between this basic technologies and the software framework. Main basic technologies are:

Object recognition, Event recognition and scene Interpretation by Technical University Vienna [4]: A system to generate hypotheses on the current state and events happening in human robot collaborative scenarios (HRC) is being developed. The software modules will be based on existing approaches and software libraries for object modelling and object pose recognition, concepts to describe events in HRC scenes, and fusion of data streams including action recognition, robot states and object recognition.

Mixed Reality methods by Evolaris [5]: Focus of this work is to develop methods to augment visual information using Head Mounted Displays (HMDs) and modes for the user to interaction with the HMD (data input). A major challenge is given through the requirement of selecting appropriate information given the current context and individual needs of the user.

Visualization of complex data by Fraunhofer Austria [6]: The main focus is developing approaches to enable real-time visualization of large amount of data, e.g. complex CAD models, on thin clients (data glasses). Moreover, a model-based tracking approach based on CAD data is developed, to facilitate position-stable augmentation of data in industrial environments.

Interaction for robot-based Assembly processes by PROFACTOR [7]: Within this technology package, concepts to enable intuitive interaction in HRC scenarios will be developed. Major challenges include the implementation of flexible models to enable fast adaptation of process knowledge and adaptation of the human-robot interaction (user specific needs), avoiding explicit programming.

Acoustic Interaction by Joanneum research [8]: The main goal is to develop speech-interfaces to enable intuitive interaction with assistance systems in an industrial setting. In order to maximize user friendliness, the interfaces are not restricted to a collection of commands and can cope with different dialects and languages. Acoustic feedback is used to inform the user about the states of the assistance system.

Iterative Interaction Design by AIT [9] and PLUS [10]: The goal is to implement a Research through Design (RtD) based process, where prototypes for current and present interaction models/modes are developed by potential end-users. This generated, valuable feedback serves as input to an iterative development process for assistance system interaction design.

IV. FIRST RESULTS

As the project has started in Mai 2017, the work performed in the first 6 months was focused on requirements and finding a set of basic technologies as described in chapter IV. Also a more detailed definition of the use cases and the Assistance Units, which will be implemented, was done. This led to 3 different Use Cases with 7 Assistance Units in total.

- **Service and maintenance (USE CASE 1)**
  - Notification of maintenance protocols (Unit 1)
  - Communication with Experts (Unit 2)
- **Setup and multi machine service (USE CASE 2)**
  - Guiding through setup process (Unit 3)
  - Multi machine service (Unit 4)
- **Assembly (USE CASE 3)**
  - Notification of Assembly instructions (Unit 5)
  - Part delivery (Unit 6)
  - Assembly instructions review (Unit 7)

ACKNOWLEDGMENT

This work has been supported by the Austrian Research Promotion Agency in the program "Production of the Future" funded project MMAssist II (FFG No.: 858623) and the Austrian Ministry for Transport, Innovation and Technology (bmvit).

REFERENCES

[1] (2018) MMAssist_II. [Online]. Available: http://www.mmassist.at
[2] (2018) MMAssist_II scientific partners. [Online]. Available: http://www.mmassist.at/partner/wissenschaftliche-partner/
[3] (2018) MMAssist_II economic partners. [Online]. Available: http://www.mmassist.at/partner/unternehmenspartner/
[4] (2018) Technical University Vienna. [Online]. Available: https://www.acin.tuwien.ac.at/en/
[5] (2018) Evolaris. [Online]. Available: https://www.evolaris.net/en/

[6] (2018) Fraunhofer Austria. [Online]. Available: https://www.fraunhofer.at/en.html
[7] (2018) PROFACTOR GmbH. [Online]. Available: http://pointclouds.org/abou
[8] (2018) Joanneum Research. [Online]. Available: ttps://www.joanneum.at/en/
[9] (2018) AIT. [Online]. Available: https://www.ait.ac.at/en/about-the-ait/center/center-for-technology-experience/
[10] (2018) PULS. [Online]. Available: https://www.uni-salzburg.at/index.php?id=52&L=1/

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

FlexRoP - Flexible, Assistive Robots for Customized Production

Markus Ikeda[1]*, Srinivas Maddukuri[1], Michael Hofmann[1], Andreas Pichler[1], Xiang Zhang[2], Athanasios Polydoros[2], Justus Piater[2], Klemens Winkler[3], Klaus Brenner[3], Ioan Harton[4] and Uwe Neugebauer[4]

***Abstract*— Flexible production assistants of the future are required to be skillful, universally applicable, safe and easy to program. State of the art robot systems that are intended to be used for human robot collaboration require in some cases unintuitive text based programming, and remain, especially in combination with peripheral hardware like external sensors or machine vision algorithms, complicated. The FlexRoP project tries to overcome current limitations by development and usage of a flexible skill-based robot programming middleware and improved user interface technologies. This paper introduces usecases, the intended system architecture, methodology for description and training of kinesthetic skills as well as first application results and intentions for future developments.**

I. Introduction

Medium to small batch size production often can't be automated with robots which require costly space and need infrastructure (e.g. fences and fixtures for part allocation). Uncertainty handling (e.g. objects that are not allocated in a defined way or underlie a tolerance in type, shape or color) is far from trivial. Additional sensors and algorithms increase system complexity and require special engineering knowledge. Flexibility for industry means universal applicability and deployment to unmodified human workplaces as far as tools or processes are concerned without complex recertification procedures or questioning legal security. Ramp up of new and recommissioning of former applications is required to be done fast and by non experts.

The FlexRoP project will carry out research to make robots easier to program and thus more flexible. Project goals comprise the definition of a universal skill representation for assembly tasks, implementation of automatic and semiautomatic skill acquisition techniques based on observation learning and kinesthetic teaching, generalization techniques and implementation of skill based action synthesis algorithms.

This paper presents:

- Two selected real world usecases in the FlexRoP project.
- The system architecture for the flexible robotic assembly assistant providing workflow based programming.
- A methodology to describe and acquire kinesthetic skills from kinesthetic demonstration.
- Evaluation results from workflow based programming with kinesthetic parameterization and kinesthetic skill acquisition.
- Inferred intentions for future developments.

[1] Profactor GmbH, Im Stadtgut A2, A-4407 Steyr-Gleink, Austria
[2] Universität Innsbruck, Innrain 52, A-6020 Innsbruck, Austria
[3] Schmachtl GmbH, Pummererstrasse 36, A-4020 Linz, Austria
[4] Magna Steyr Fahrzeugtechnik AG & Co KG , Liebenauer Hauptstrasse 317, A-8041 Graz, Austria
*Corresponding author: Markus Ikeda; Markus.Ikeda@profactor.at

II. Usecases

Two real world production usecases from automotive preassembly are considered. The usecases require screwing, clip in and manipulation operations in a very broad range of applications. In so called brownfield [1] environments available (hand)tools have to be picked up by the robot rather than spanning specialized robot tools to guarantee deployability to any human workplace.

Usecase A targets the pre-assembly of a centerspeaker assembly. A speaker has to be fixed with three screws to a plastic carrier while a tweeter needs to be clipped in (see Fig. 1). Handling of the non-rigid wires is omitted. Process forces are low but the required pose precision for screwing and clipping is very high ($<$ 1mm). The complexity of the entire process (which consists of 7 subprocesses - see Table I) is extremely high. Three different objects are presented in boxes and have to be manipulated as well as the intermediate assemblies and the power screwdriver which has to pick up, hold and manipulate the screw axially perfectly aligned during transport and process. In order to be able to guarantee product and process quality methodology for quality assessment is required. This might be natural and easy for a human but independent of the available data (acoustic, FT-signal, optical) extremely challenging for any technical system.

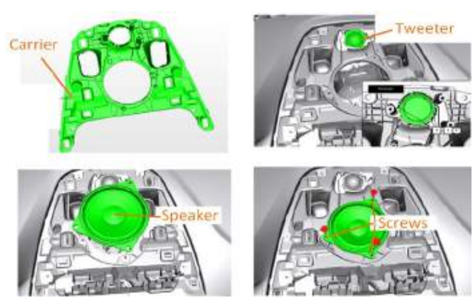


Fig. 1. Usecase A - Center-speaker assembly

Usecase B considers the joint pre-assembly of an automotive swivel-bearing assembly by human and robot. Han-

dling of components and assembly takes place close to the robot's load limits with high handling and process forces. A human carries out processes unsuitable for the robot like screw feeding and delicate ambidextrous assembly operations (e.g. mounting of brackets and brake hose - 1 in Fig. 2). Unergonomic handling of heavy objects is carried out by the robot as well as the error prone screw tightening operation for the assembly of wheel bearing to swivel bearing which needs to be carried out in a specific order (2 in Fig. 2).

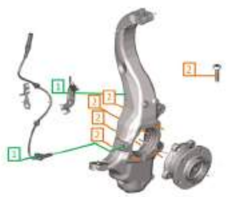


Fig. 2. Usecase B - Swivel bearing assembly

III. System Architecture

A. Hardware

The robot assistant (Fig. 3) consists of a passively mobile platform with retractable wheels, an electric enclosure containing robot and system controller as well as additional IO and power supply components. The platform is equipped with a KUKA LBR iiwa 14 R820 robot. The User Interface (UI) consists of a touch screen monitor on the mobile platform and the robot's touch pneumatic media flange. The robot is equipped with one universal tool for both applications. The diversity of requirements with regard to object shapes, payloads and processes result in a complex tool design (see Fig. 4) with following components:

- Force Torque (FT) sensor for measuring process wrench.
- A chassis for installation of various components.
- RGBD and 2D cameras for automatic position accuracy compensation functionality.
- Two electric grippers in order to be able to manipulate multiple objects or long objects.
- Automatic toolchanger for spanning additional process tools (ordinary hand tools articulated by pneumatic actuators).

Handling and manipulation of objects was intendend with universal grippers and force closure. Tests disproved the applicability of several universal grippers for accuracy and process stability reasons so aluminium fingers with form adjusted plastic inlays are used.

B. Software

The robot assistant is required to be programmable without special training. A KUKA iiwa [2] may as a HRC-capable device offer handguidance for parameterization but needs to be programmed text based (in JAVA) as well as

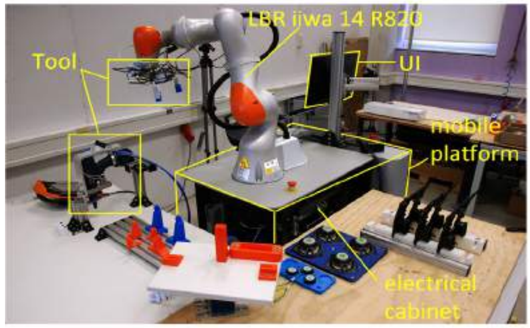


Fig. 3. System Overview

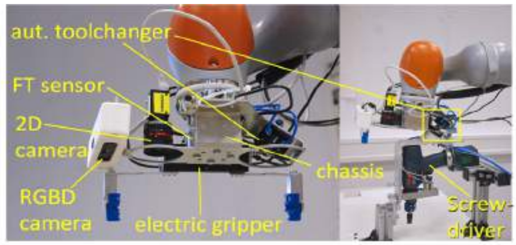


Fig. 4. Flexible tool prototype (2nd gripper not installed)

machine vision algorithms or standard PLC code. Therefore XRob™ [3] (see Fig. 6) is introduced as an abstraction layer for all hardware (cameras, sensors, robots, etc.) and software components (object pose recognition, path planning, etc.). For kinesthetic skill learning a real time interface to the robot and the FT sensor is required. Therefore ROS and the KUKA fast research interface are used. Fig. 5 describes the selected modular system architecture.

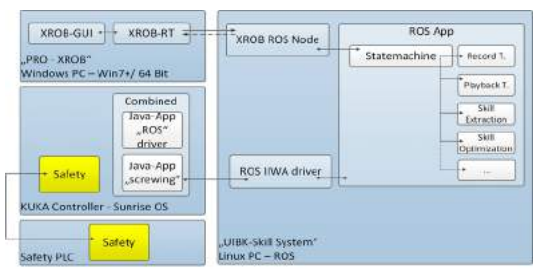


Fig. 5. Software architecture

IV. Skill Based Workflow Programming

Robot programming in industrial applications is done mainly in proprietary text based programming languages. Skills are treated as traditional, "unintelligent" robot motion programs (macros) that are augmented with pre and

post-conditions to add situational knowledge. Macros are supposed to work on objects in the workspace that are recognized via some kind of sensing device (e.g. optical). For example [4] presents a unifying terminology for task-level programming of highly flexible mobile manipulators in industrial environments, while [5] demonstrates the skills which are needed for industrial kitting applications.

A. Skill Based Programming Framework

Task-level programming is based on lower level entities, usually called robot skills. The description of processes can be done at different levels of granularity. Tasks can be broken into more or less complex subtasks ranging from sensory and/or motor base skills to complex aggregate subtasks. A skill is a primitive that allows the coordination, control and supervision of a specific task. The primitives can incorporate advanced task specifications, necessary control, and sensing capabilities, which allows a skill to handle uncertainties during execution. In contrast to the concept of skills, skill primitives [6], [7], [8] are rather well defined in the robotics community. This layered approach is reflected by the design of the XRob^{TM}software framework (see Fig. 6) which can aggregate basic functionality (e.g. data acquisition, image processing, robot movements and macros, etc.) to more complex aggregate subtasks that can easily be reused. After graphical configuration of a workflow process points are parameterized by bringing the tool center point to its destination and adopting relevant data (e.g. the current position or the current camera image) electronically. That allows programming processes and movements between quasistatic intermediate process points. If more complex trajectories are required the system incorporates dynamic motion primitive based skills.

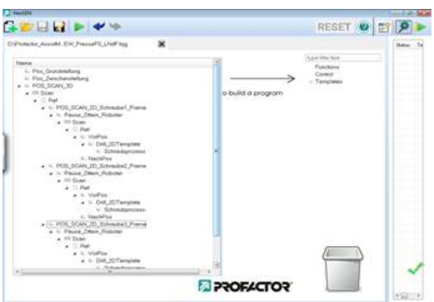


Fig. 6. XROB Graphical User Interface

B. Dynamic Motion Primitive Based Skills

In [9] and [10] is given an overview on programming by demonstration. Dynamic Motion Primitives (DMPs) have been a very popular method for learning and generalization of kinesthetically taught motions [10] with multiple extensions [11], [12], [13]. They are motivated by the need to derive a motion representation which is capable, not only to reproduce complex trajectories but also to easily generalize them. DMPs are a combination of two terms. A simple linear dynamical system $l(\cdot)$, which is well defined and has stable behavior and a nonlinear forcing term $f(\cdot)$ which makes the reproduction and generalization of complex motions feasible

$$\ddot{y} = l(g, y, \dot{y}) + f(x, g). \tag{1}$$

In the case of discrete motions the linear system is a stable attractor, usually a PID controller

$$l(g, y, \dot{y}) = \alpha_y(\beta_y(g - y) - \dot{y}), \tag{2}$$

where y is the joints' position of the robot, g is the target states, and α and β are gain terms of the PID controller which draw the manipulator to the target state. Adding a forcing term to the linear system allow to modify the trajectory:

$$\ddot{y} = \alpha_y(\beta_y(g - y) - \dot{y}) + f. \tag{3}$$

The challenge in DMPs is to appropriately define the non-linear forcing term f over time while ensuring stability of the system and generalization. This is achieved by introducing a canonical dynamical system denoted as x with simplistic dynamics:

$$\dot{x} = -\alpha_x x. \tag{4}$$

Thus the forcing term f depends on the value of the canonical system as follows:

$$f(x, g) = \frac{\Sigma_{i=1}^{N} \psi_i w_i}{\Sigma_{i=1}^{N} \psi_i} x(g - y_0). \tag{5}$$

y_0 is the starting state of the system, and $\psi_i = \exp\left(-h_i\left(x - c_i\right)^2\right)$ is a Gaussian kernel centered at c_i.

Training of DMPs is achieved by optimizing its hyper-parameters (w) with a given trajectory. While the desired motion is demonstrated, the sensors' values are recorded and they are used to derive the hyper-parameters based on Eqn. (3) which is written as:

$$\ddot{y} - \alpha_y(\beta_y(g - y) - \dot{y}) = f(x). \tag{6}$$

Thus, the forcing term is optimized to compensate the error of the linear dynamical system – which are the training targets of the learning rule – at each state of the canonical system x which is the training input. This corresponds to a regression problem which can be solved with a variety of methods such as Locally Weighted Regression [14] or Locally Weighted Projection Regression [15].

C. Motion Assessment Primitive

The motion assessment primitive is responsible for providing an evaluation of the performed motion, thus it evaluates the DMP's performance. This is achieved by a two-tier process which exploits the trajectory recorded through kinesthetic teaching. Those recordings include both the joints' states and the exerted force/torques on the end-effector. This makes the derivation of the motion's contact dynamics

model through machine learning techniques feasible which maps joint states to exerted forces/torques. Thus the system "learns" which forces and torques to expect at specific joint states. Therefore, a ground-truth model is created from the end-user demonstration and a comparison model is created from the recording of the autonomous DMP's movement reproduction. The difference of those two models is measured and fed to the second tier which classifies the motion as successful or not.

Gaussian processes (GPs) are employed for learning the wrench model of the executed task. GPs are a powerful non-parametric machine learning approach. Contrary to other methods that infer a set of function parameters, GP infers the function f directly and therefore can be anticipated as probability distribution over functions. A GP is defined by a mean $m(\mathbf{x})$ and a kernel (covariance function) $\mathbf{K}(\mathbf{x},\mathbf{x})$ as illustrated in Eqn. (7). Typical choices are a squared exponential kernel and a zero mean

$$f(x) \sim \mathcal{GP}(m(\mathbf{x}), \mathbf{K}(\mathbf{x},\mathbf{x})). \tag{7}$$

GP, employ the Bayes rule for the derivation of the posterior distribution over functions -see Eqn. (8), where $\mathbf{t}$ is the vector of target values, the force/torques in this case. In regression problems the latent function f is continuous and therefore an appropriate likelihood is the normal distribution $\mathcal{N}(f|\mathbf{m}(\mathbf{x}), \mathbf{K}(\mathbf{x},\mathbf{x}))$ and the GP prior is also a Gaussian process $p\left(f|\mathbf{X}\right) \sim \mathcal{GP}\left(0, \mathbf{K}(\mathbf{x},\mathbf{x})\right)$.

The posterior distribution which represents the learned wrench model given the recorded data .

$$p\left(f|\mathbf{X},\mathbf{t}\right) = \frac{\mathcal{N}(f|\mathbf{m}(\mathbf{x}), \mathbf{K}(\mathbf{x},\mathbf{x}))p\left(f|\mathbf{X}\right)}{p(\mathbf{t}|\mathbf{x})}. \tag{8}$$

The term of interest in the case of motion assessment primitive is the marginal likelihood $p(\mathbf{t}|\mathbf{x})$ because the optimal parameters of the kernel are derived by optimizing it. Thus, the contact dynamics model derives by minimizing the logarithm in Eqn. (9) which can be achieved by using any gradient-based optimization method such as gradient descent

$$\log(\mathbf{t}|\mathbf{X}) = -\frac{1}{2}\mathbf{t}^T\mathbf{K}^{-1}\mathbf{t} - \frac{1}{2}\log|\mathbf{K}| - \frac{n}{2}\log 2\pi. \tag{9}$$

The assessment primitive creates six ground-truth models, one for each wrench degree of freedom and other six models from the autonomous execution of the DMP. Those probabilistic models are then compared using Hellinger distance - Eqn. (10) - which yields a similarity measurement h for each k wrench component. Those measurements are fed to the second tier of the primitive, a Naive Bayes classifier which classifies the similarity measures as success or failure

$$h_k\left(\mathcal{GP}_k^{\text{demo}}, \mathcal{GP}_k^{\text{exe}}\right) = \sqrt{1 - \frac{\sqrt[4]{|\boldsymbol{K}_k^{\text{demo}}||\boldsymbol{K}_k^{\text{exe}}|}}{\sqrt{\frac{1}{2}|\boldsymbol{K}_k^{\text{demo}} + \boldsymbol{K}_k^{\text{exe}}|}}}. \tag{10}$$

In the second stage of the primitive the set of similarity measurements $\mathbf{h}$ are fed to a Naive Bayes classifier which applies the Bayes rule - Eqn. (11) - for the derivation of $p(C_j|\mathbf{s}^\star)$, where $p(C_j) = N_j/N$ is the prior probability of the class j, $p(\mathbf{h}^\star|C_j)$ is the likelihood that the sample $\mathbf{h}^\star$ belongs to the class j and $p(\mathbf{h}^\star)$ is a scaling term independent from the class and therefore can be omitted

$$p(C_j|\mathbf{s}^\star) = \frac{p(C_j)p(\mathbf{s}^\star|C_j)}{p(\mathbf{s}^\star)}. \tag{11}$$

The likelihood derives based on the assumption that the similarity measurements are independent and identically distributed and is calculated as:

$$p(\mathbf{h}^\star|C_j) = \prod_{d=1}^{D} p(s_d^\star|C_j), \tag{12}$$

where K is the dimensionality of the similarity measurements. It is assumed that their values are distributed according to a Gaussian distribution $\mathcal{N}\left(\mu_k^j, \sigma_k^j\right)$ with mean μ_k^j, the mean value of similarity measurement d which belongs to class j and its corresponding variance σ_k^j. Thus Eqn. (11) can be written as:

$$p(C_j|\mathbf{s}^\star) \propto p(C_j)\prod_{d=1}^{D}\mathcal{N}(s_d^\star|\mu_d^j, \sigma_d^j), \tag{13}$$

where the parameters of the Gaussian distribution derive by maximum likelihood estimation.

D. Intermediate Results

1) Skill Based Workflow Programming: FlexRoP identified macros for screwing operations as well as the clip-in operation that can be considered as robotic skills themselves and serve as baseline for performance comparison with the kinesthetic skills developed by the project.

The screwing macro considers the basic parameters: start pose, screw length and process force. The clipping macro considers in a similar way start pose, end pose and process force. Together with parameterizeable macros for other operations (robotic movements,etc.) screwing and clipping are accessible through XRob^{TM}.



Fig. 7. Overview - Usecase A

Usecase A was split into several suboperations and program templates were created accordingly. Parameterization

of the templates was done by moving the robot to a specific process point and recording relevant data (e.g. cartesian positions, reference images, gripper opening,...). For usability reasons movement of the robot was planned to be done by hand guidance. Tool parameters (inertia, mass) are tuned and the robot flange is intended to be used in zero gravity mode. Total parameterization time of existing workflows sums up to 285min (see Table I), which is high (compared to a target time of 30min which is derived from a productivity calculation) and was caused by high accuracy demands to be able to achieve process stability. Tight clearances of carrier plate and fixtures as well as required positioning accuracy for screws ins screw-holes and components to be assembled require precise teach in which cannot be achieved in gravity compensation mode. For perfect vertical tool alignment and fine positioning of the tool it was, due to not available interfaces required to use the robot teach pendant which required several stop and start operations of the XRob™driver on the robot controller as well as operation mode changes from automatic to hand mode and vice versa in order to be able to use the robot teach pendants integrated positioning utilities. A detailed analysis of subprocess 1 (see Fig. 8) reveals that operation mode changes as well as interaction with the GUI of the robot (which is required to select correct coordinate frames to travel in for fine positioning or selection of speed) in addition with finepositioning itself is accountable for almost two thirds of the reparameterization time. Interaction with the XRob™(HTML-)GUI and adjustment of the finger positions in comparison requires less time.

TABLE I

AVERAGE PARAMETERIZATION TIME - 3 TRIALS

subprocess	description	parameterization time
1	Move carrier from rack to assembly fixture	30min
2	Move speaker from rack to assembly fixture (via rotation table)	60min
3	Pick-up of power tool from pod	15min
4	Screw-pick-up & screwing operation (three target positions)	45min
5	Deposit of power tool to pod	15min
6	Reorientation of assembly	60min
7	Move tweeter from rack to clip in position and clipping operation	60min

2) Dynamic Motion Primitive Based Skills: In order to evaluate the performance of both the motion and assessment primitives a mock-up which imitates the project's clip-in process was designed (see Fig. 9). For the evaluation a KUKA iiwa equipped with an ATI force/torque sensor and a simplistic suction cup was used. The primitives were trained on recorded data from one single kinesthetic demonstration and their generalization ability is tested by varying the start pose of the manipulator. The motion primitive managed successfully to execute 17 out of 44 trials resulting in a 39% success rate. An illustration of a successful snap is presented in Fig. 10. The motion assessment primitive was evaluated off-line on datasets collected from 31 motions using a cross validation method for training and evaluating the Naive Bayes classifier. In this evaluation method, the data is partitioned in training and testing datasets. The former are used for optimizing the hyper-parameters of the Naive Bayes classifier while the later for evaluating its performance. In detail, a leave-one-out cross validation is performed where the classifier is trained with all the datasets except one which is used for testing. This iterative procedure finishes when all the datasets have been used for testing.

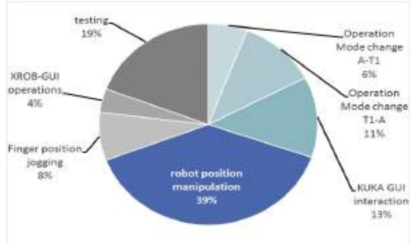


Fig. 8. Subprocess 1 - teach in time breakdown



Fig. 9. The snap-fit used for performing evaluation of the motion and assessment primitives.

V. CONCLUSIONS & FUTURE WORK

Two immediate directions for further improvements were identified.

A. Skill Based Workflow Programming

Experiments Showed that workflow based programming is still complicated for untrained users. Programming in the worker's domain without kinesthetic manipulation of the robot itself remains desirable. A novel instrumented power tool as UI for teach in operations is planned. A worker will not have to specify numeric values on a GUI in order to parameterize process workflows. The instrumented power tool will record trajectories in 6DOF as well as time series of process forces and torques as well as the actuation of the tool. Startposition, process forces and screw lengths will



Fig. 10. KUKA iiwa performs a successful snap-fit using DMPs

be derived from the analysis of the data. In comparison to kinesthetic teach in the so called embodiment problem has to be solved since the robot has different reach and multiple kinematic configurations that can be used to position a tool.



Fig. 11. Instrumented tool concept

B. *Dynamic Motion Primitive Based Skills*

The future work on DMPs will be focused on the issue of the low success rate. A reason for the low performance could be that DMPs create a single model for each degree of freedom. Valuable information regarding the correlations which exists between the joints' states an the exerted forces/torques my be lost. This can be dealt with using multi-modal motion representations which couple the joint state with the exerted forces/torques and thus create a single model using all the sensory inputs.

Furthermore, motion assessment is currently performed after the completion of the motion. A possible expansion is to assess the motion during runtime. This would significantly decrease the chance of damage for both the robot and the manipulated object. A minor issue is the high computational complexity of GPs which affects the time needed for assessment, especially on long motions. Therefore, it is planned to investigate the applicability of other, more computationally efficient models.

Finally, the main focus of the future work will be given on the development of a motion optimization primitive. This would optimize the hyper-parameters of the DMPs in such a way that the probability of a successful motion is maximized and thus will close the loop between motion and assessment primitives. The machine learning approach which will be used belongs to the class of reinforcement learning. In detail, the contact dynamics model could be exploited and so the DMPs will be optimized based on simulations of the learned model instead of the real system. Such an approach belongs to the class of model-based reinforcement [16] learning which has advantages such as minimal optimization time and also minimal risk of damage for both the robot and the manipulated objects which makes it appropriate for manufacturing tasks.

ACKNOWLEDGMENT

FlexRoP is funded by the Austrian Federal Ministry of Transport, Innovation and Technology (BMVIT) under the program "ICT of the Future" between [2016-09 and 2018-08]. More information `https://iktderzukunft.at/`.

This project is supported by grant funding "FTI Struktur Land Oberšterreich".

This project is supported by grant funding "IWB/EFRE-Programm Österreich 2014-2020".

REFERENCES

[1] R. Hopkins and K. Jenkins, *Eating the IT Elephant: Moving from Greenfield Development to Brownfield*, ser. IBM Press. Pearson Education, 2008.

[2] (2018) KUKA IIWA. [Online]. Available: http://www.kuka-robotics.com/en/products/industrial_robots/sensitiv/lbr_iiwa_7_r800/start.html

[3] S. C. Akkaladevi, M. Plasch, and A. Pichler, "Skill-based learning of an assembly process," *e & i Elektrotechnik und Informationstechnik*, vol. 134, no. 6, pp. 312–315, Sep 2017.

[4] F. Rovida, M. Crosby, D. Holz, A. S. Polydoros, B. Großmann, R. P. A. Petrick, and V. Krüger, *SkiROS—A Skill-Based Robot Control Platform on Top of ROS*. Cham: Springer International Publishing, 2017, pp. 121–160.

[5] A. S. Polydoros, B. Großmann, F. Rovida, L. Nalpantidis, and V. Krüger, "Accurate and versatile automation of industrial kitting operations with skiros," in *Towards Autonomous Robotic Systems*, L. Alboul, D. Damian, and J. M. Aitken, Eds. Cham: Springer International Publishing, 2016.

[6] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in Cognitive Sciences*, vol. 3, no. 6, pp. 233 – 242, 1999.

[7] A. Björkelund, L. Edström, M. Haage, J. Malec, K. Nilsson, P. Nugues, S. G. Robertz, D. Störkle, A. Blomdell, R. Johansson, *et al.*, "On the integration of skilled robot motions for productivity in manufacturing," in *Assembly and Manufacturing (ISAM), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1–9.

[8] N. Krüger, J. Piater, F. Wörgötter, C. Geib, R. Petrick, M. Steedman, A. Ude, T. Asfour, D. Kraft, D. Omrcen, *et al.*, "A formal definition of object-action complexes and examples at different levels of the processing hierarchy," 2009.

[9] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 2, pp. 286–298, 2007.

[10] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469 – 483, 2009.

[11] M. Deniša, A. Gams, A. Ude, and T. Petrič, "Learning compliant movement primitives through demonstration and statistical generalization," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 5, pp. 2581–2594, 2016.

[12] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in neural information processing systems*, 2013, pp. 2616–2624.

[13] D. H. Park, H. Hoffmann, P. Pastor, and S. Schaal, "Movement reproduction and obstacle avoidance with dynamic movement primitives and potential fields," in *Humanoids 2008 - 8th IEEE-RAS International Conference on Humanoid Robots*, 2008, pp. 91–98.

[14] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning for control," *Artificial Intelligence Review*, vol. 11, no. 1, pp. 75–113, Feb 1997.

[15] S. Schaal, C. G. Atkeson, and S. Vijayakumar, "Scalable techniques from nonparametric statistics for real time robot learning," *Applied Intelligence*, vol. 17, no. 1, pp. 49–60, Jul 2002.

[16] A. S. Polydoros and L. Nalpantidis, "Survey of model-based reinforcement learning: Applications on robotics," *Journal of Intelligent & Robotic Systems*, vol. 86, no. 2, pp. 153–173, May 2017.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Development of a 3D-Printed Bionic Hand with Muscle- and Force Control

Florian Dannereder[1], Paul Herwig Pachschwöll[1], Mohamed Aburaia[2], Erich Markl[2], Maximilian Lackner[2], Corinna Engelhardt-Nowitzki[2] and Diane Shooman[2]

***Abstract*— The majority of people with upper extremity loss replace their arm and hand with a low-cost prosthesis. However, an average prosthesis only covers minimal functionality in comparison to a human hand, and the user is strongly limited in everyday life. Sophisticated bionic hands have been developed to replace upper extremity functionality. A bionic hand can be controlled via muscle contraction of the upper extremity or the shoulder area, and can replace the main functions that a human needs in everyday life. Nearly every hand movement and the independent movement of the fingers can be produced through a rotation mechanism around the wearer's wrist. Since these bionic hands are very expensive, only a small percentage of the world population have the privilege to own one. To close the gap between customer, designer and engineer, an open source bionic hand that can be 3D-printed is a cost effective possibility. The result of this project is a cost effective 3D-printed bionic hand that can be reprogrammed for user specific functions. The sensed muscle regions can be changed spontaneously as needed. The sensitivity of the muscle contraction and the gripping force are adjusted by software using a closed loop control.**

I. INTRODUCTION

Mastering the use of a bionic hand to manipulate objects in our daily environment can be so complex, that numerous of users revert back to simpler prosthetics. A particular technical challenge in bionic hand design is to create an effective interface for the wearer, and to provide a wide spectrum of grip types through muscle control. Individual differences in each human body influence the control algorithms and the muscle contraction detection. To improve the daily use, some personal settings e.g. different speeds, thresholds or grips should be adjustable. This paper describes a 3D-printed bionic hand with 15 different gripping styles, which can be controlled by muscle contraction from the upper extremity. It provides an automatic stop of the finger movement when touching an object at a determined force, although the users muscle is still contracted. This simplifies the bionic hand control through muscle contraction and has a tremendous impact on controllability.

II. STATE OF THE ART

Modern bionic hands are controlled by myoelectric signals, which allow precise control of different grips. Those myoelectric signals sense a chosen muscle region that is contracted by the prosthesis user. With this method the amputees brain is capable of controlling the bionic hand with good accuracy and low difficulty [1]. Currently, the most popular commercial bionic prostheses with high-technical functionality are the Touch Bionics I-Limb-Ultra and the Bebionics RSL Steeper. This two bionic hands, shown in Figure 1, will be discussed in this chapter.

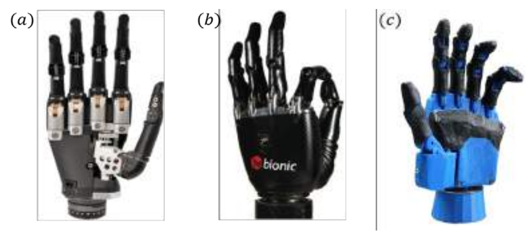


Fig. 1. High technical bionic hands (a) Touchbionics iLimb Ultra [2], (b) RSL Steeper Bebionic [3]

High technical bionic hands, which can replace upper extremity functionality, are realized with eleven joints. In comparison, a real human hand has 33 joints [4]. The most bionic hands facilitate a finger movement with a coupler mechanism, or with a tendon linkage. The eleventh joint is the thumb slewing mechanism, to change between an open and a closed hand. Different finger mechanisms with up to two joints are shown in Figure 2.

An important characteristic of the finger construction is a self-locking mechanism, which can be carried out in different ways. A self-locking mechanism is important for the end positions of the fingers, to prevent some inadvertent position change of the current activated grip. The I-Limb Ultra uses a DC motor with a spur gear, to transfer the torque to a worm gear. In contrast, the RSL Steeper uses a linear DC motor with an integrated lead screw. This makes it possible for both bionic hands to block the finger-movement while the motors are turned off. In Table I different technical specifications are shown.

III. PROBLEMS AND CHANCES

The typical muscles of a prosthesis wearers forearm are not always useable, which means that the position of the electrodes has to be selectable. In case of only one useable forearm muscle, a shoulder or an upper arm muscle can

[1]Florian Dannereder and Paul Herwig Pachschwöll are students of the study program Mechatronics/Robotics at the University of Applied Sciences Technikum Wien, Austria {florian.dannereder, paul_herwig.pachschwoell}@technikum-wien.at

[2]Mohamed Aburaia, Erich Markl, Maximilian Lackner, Corinna Engelhardt-Nowitzki and Diane Shooman are with the department of Advanced Engineering Technologies at the University of Applied Sciences Technikum Wien, Austria {aburaia, markl, lacknerm, engelhac, shooman}@technikum-wien.at

TABLE I
TECHNICAL SPECIFICATIONS OF THE I-LIMB ULTRA AND THE RSL STEEPER [5]

Product	I Limb Ultra	RSL Steeper
Developer	Touch Bionics	Otto Bock
Weight	405-479g	495-539g
Number of Joints	11	11
Number of Actuators	5+1 (motorized thumb)	5
Actuation Method	DC Motor-Worm Gear	Linear DC Motor-Lead Screw
Joint Coupling Method	Tendon Coupling	Coupler mechanism

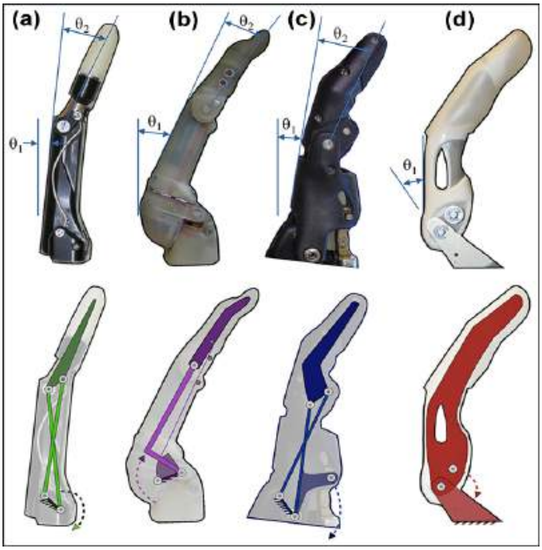


Fig. 2. Different finger mechanisms used for bionic hands with up to two degrees of freedom, (a) Vincent, (b) I-Limb, (c) RSL Steeper, (d) Michelangelo [5]

be used instead. The orientation of the bionic hand before picking up an object is also important. The only way to do so is through a rotation made by the bionic hands wrist, which adds another degree of freedom. This function is used to tilt a bottle and fill a cup. To achieve a tight grip on the bottle, every finger is powered by an actuator, which allows an independent finger movement. To switch between an open hand and a grip for taking a bottle, a slewing thumb is necessary. The most common grips do not always need every finger, therefore some precision grips have been programmed. Picking up a pencil from a table can be done with the use of three fingers. To avoid a vibrating and noisy bionic hand, self-locking actuators have been combined with a coupler mechanism. The bionic hands actuators create a high force, which create the necessity of a force control when the hand closes, to avoid damaging itself or the objects gripped by the fingers.

A. Placement of the Muscle Sensors

Interpretation of sensed muscle contraction is a complex procedure. Somehow, when a muscle is contracted, the opposite muscle contracts softly too. This creates the possibility of using different electrode modes, like a single mode or a dual mode. Raw detected muscle signals with a total of 105 measurement points on the x-axis, recorded in a time of two seconds is shown in Figure 3. The y-axis represents the 10-bit ADC-value from the Myoware muscle sensors. For the data point evaluation, a threshold has to be set, which defines whether the muscle was seriously contracted or not. If a digital value of more than 160 were interpreted as a positive muscle contraction, every signal jump exceeding a threshold of 160 would be read as a single detected muscle impulse. This makes toggling functions by using muscle contraction difficult. The measurement points 92, 97 and 102 show fast signal jumps, with a severe influence on the muscle contraction detection [6]. These three short impulses are the result of contracting the opposite muscle region.

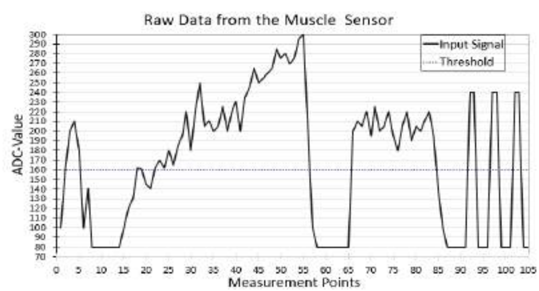


Fig. 3. Recorded signal of two short muscle contractions of the forearm, converted into digital values

To avoid a false detection, a filter is the best solution, nevertheless it is possible to attach the electrode to another muscle region. Another common muscle region is the shoulder area, which cannot be influenced by the forearm muscles. Concerning the fact that almost every muscle can be detected, it is possible to connect the bionic hand with another muscle electrode. A threshold of an ADC-value of more than 160 would create the following interpretation of the muscle activity.

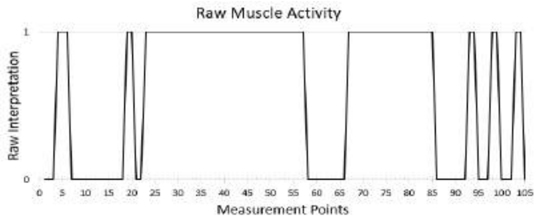


Fig. 4. Using a threshold of 160 to decide if a muscle is contracted or not (0 = relaxed, 1 = contracted)

B. Independent Finger and Wrist Movement

Without a rotating wrist, the shoulder joint would be the only way to adjust the orientation of the bionic hand. Small jobs like filling a liquid into a cup would become very difficult without an additional degree of freedom. Grabbing a bottle needs an encircling thumb that ensures a steady grip [7]. A bionic hand with an independent finger movement can be used for many different grips, which makes it very useful in everyday scenarios. The coupler mechanism makes it possible to transform the actuators linear movement into a finger movement that keeps the relation between its travelled distance and the position of the fingers. The necessary force to close the hand is not constant and changes while closing. A measurement of the idle closing current of each finger involves further details. The actuator rod travels 13.5mm to convert the open palm into a closed palm with every finger in its bend position, which is explained in Figure 10. This delivers an amount of measurable steps for the force control, which monitors the force by finger movement. The current is not steady, which makes the use of a simple current limit inaccurate. A method had to be developed to use the collected current data and create a more accurate force control. Nevertheless, each finger has its own current progress, so the created method has to be flexible.

IV. METHODS

The 3D-printed bionic hand is an open-source device made only of nonindustrial components to ensure that every interested person is able to reconstruct it. The microcontroller is an Arduino Uno, which realises 15 different gripping styles. The movement of the fingers is made with linear actuators, which create an independent finger movement. To keep a human-like shape the bionic hand has five fingers, a rotating wrist and a pivoting thumb. The prosthesis is controlled by muscle activity and allows a high usability.

A. Implementation of a Filter for Muscle Noise Reduction

The bionic hand is designed to be controlled by two muscle electrodes. The usability as a forearm prosthesis makes using the forearm muscles the obvious choice. The placement of the electrode has an important role for the controllability. Contraction of the opposite muscle occurs spontaneously, which makes it hard to differentiate between a seriously and a spontaneously contracted muscle. The muscle signal displayed in Figure 5 shows two short muscle contractions with small signal jumps. The easiest way for smoothing those irregularities would be a first order low pass filter [8]. The drawback of using circuits is that the signal is stored in a component, which changes the speed of the signal processing significantly. Each finger would need to be equipped with such a circuit, nevertheless space inside a bionic hand is limited. Therefore, a digital solution was created. Another benefit of a digital solution is that a special behaviour can be forced. Processing the raw muscle signal from the electrodes can then be optimised to react differently on a rising signal rather than on a falling edge. This enables integration of the raw signal to create a noise reduction but which edges off the signal if a level drop is measured. The following figure shows the raw signal from Figure 3 explained in Chapter III-A, which has been filtered for better controllability of the bionic hand.

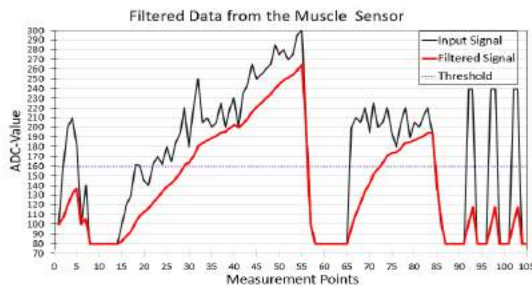


Fig. 5. Using a threshold of 160 to decide if a muscle is contracted or not (0 = relaxed, 1 = contracted)

A closer look at the figure shown above makes it clear that the similarity of the signals is still given. The digital filter was optimized to delete narrow and high jumps without deformation of the signal sequence. The mentioned edge detection allowed a fast adjustment on falling edges. The following figure shows the interpretation of the filtered signal.

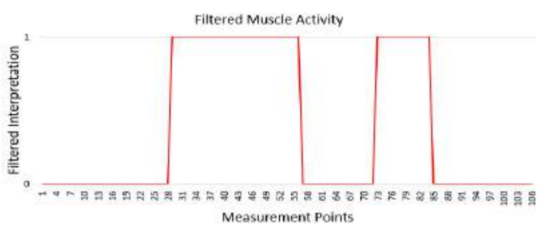


Fig. 6. Interpretation of the digitally filtered signal of two short muscle contractions by using a threshold of 160 (0 = relaxed, 1 = contracted)

Compared to Figure 4 (Chapter III-A) the difference is easily noticeable, and a better controllability of the bionic hand was achieved.

B. Estimation of the Gripping Force

Calculation of the gripping force has to be fast, applicable for every finger, and with low computing consumption. The force transfer of the actuator to the finger is a linear translation, which makes creating an accurate mathematical formula possible. A lookup table of every fingers idle current was recorded and used instead. With this individual information, the position related current limit was defined as the idle current increased by 35 percent. This method turned out to be precise enough, regardless of the position of the actuator or finger. Related to the different finger size, the relation between actuator current and finger force is individual for every finger, so the thresholds were finger-size related too. With this method the current limit shown below was defined and used for finger.

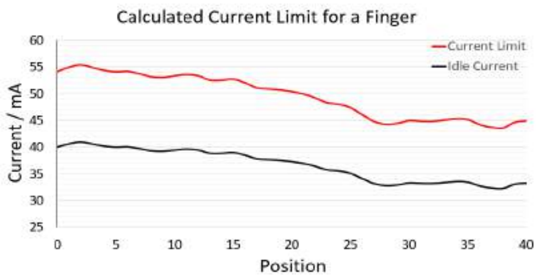


Fig. 7. Calculated current limit for a finger by using a current increase of 35 percent to create a specific current limit

Tests have confirmed that a 35 percent increase is enough to ensure a tight and reliable grip. This factor is adjustable and can be set specific for a grip.

C. Construction

As seen in Figure 8, the bionic hand has seven servomotors, which enables seven degrees of freedom. In each finger, an individual six-joint linkage is integrated, to perform a particular movement profile. The forearm is equipped with an extra motor for the wrist rotation, such as the controller and additional electronic components. In the next two sections, the realization of the finger movement and the thumb slewing mechanism will be explained in detail.

1) Finger Mechanims: For the construction, a few necessary requirements must be considered. To ensure that the bionic hand supports a two-joint finger movement, a six-joint linkage is integrated in each finger. The six-joint linkage provide a finger movement with two constrained angles θ_1 and θ_2. Therefore, a real human finger movement can be reproduced. Figure 9 shows a retracted and an extended finger position with the constrained angles θ_1 and θ_2.

The reason for using a linkage was, that it is possible to combine it with a self-locking linear servo motor. The motor is the PQ12-R micro linear servo motor from Actuonix, with a total stoke of 20mm, and a maximal linear force of 50N [9]. The linear servo motor will fully retract the motor shaft with a 2.0 ms pulse signal, and a 1.0 ms pulse signal will fully extend the motor shaft. Therefore, every position from 0mm to 20mm is approachable with the associated pulse signal [10]. The self-locking mechanism is a necessary requirement for the end positions of the fingers. The finger-linkage is constrained via a two-joint couple to the motor shaft, and therefore different finger positions by the linear movement of the motor shaft is accessible. The finger joints rotate around the instantaneous center (IC) of rotation (P_1, P_2), which are mounted into a fixed bearing inside the hand cover. The motor is also fixed inside the hand cover, and the motor shaft can move linearly.

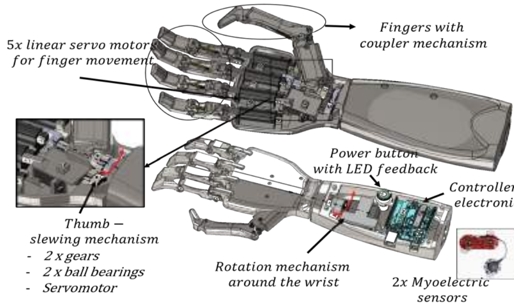


Fig. 8. CAD-Model

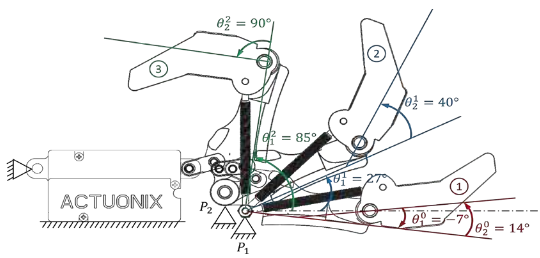


Fig. 9. Two joint finger movement with the constrained angles and , realized with a six-joint finger linkage

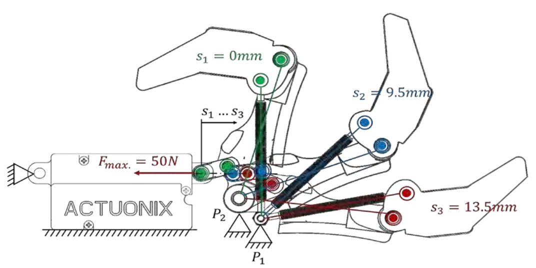


Fig. 10. Used coupler mechanism for finger movement with a total motor stroke of 13.5mm

The integrated motor type is called MG90S, and is a micro servo motor with a torque of 0.1962 Nm [11]. It is connected with a spur gear, which transfers the torque to the rotation axis of the thumb. The rotation axis of the thumb meshes with another gear, and is mounted with two ball bearings (see Figure 11).

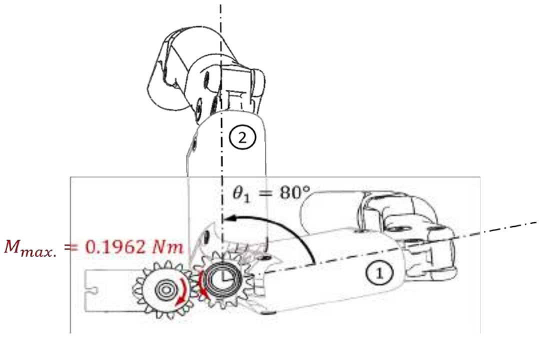


Fig. 11. Used coupler mechanism for finger movement with a total motor stroke of 13.5mm

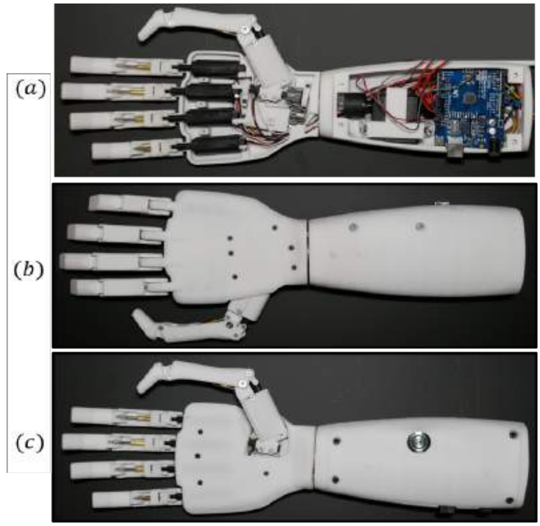


Fig. 12. Fully assembled 3D-printed bionic hand, (a) Assembled bionic hand without cover, (b) Topview internal side, (c) Topview external side

V. RESULTS

To find possible grips for programming the bionic hand, the force feedback was the first implemented function. This enabled the testing of different daily objects inside the bionic hands palm. Based on test scenarios a total set of 15 different grips has been programmed. The first tests were conducted without the use of muscle sensors, mechanical switches were used instead. Later on, the muscle sensors and its noise cancelling functions were added, which made the bionic hand controllable by muscle contraction. The way each grip is used is different, therefore the speeds for operating the fingers have been adjusted as well. The controllability of the bionic hand is precise enough to pick up a small resistor from the table surface. The used controller has still some pins left for further extensions and approximately 15 percent of its memory is available. The placement of the muscle sensors can be chosen arbitrarily, depending on the chosen muscle areas. The independent finger movement allows grips that cover up to 85 percent of the commonly used gripping scenarios.

A. A fully assembled 3D-printed bionic Hand

The final assembled bionic hand can be seen in Figure 12.

The 3D-printed bionic hand can be separated into three main assembly parts. The first assembly parts are the 3D-printed components. 3D-printed components are for example the fingers, different covers and other components designed special for this bionic hand, and are printed with a selective laser sintering-printer. In sum, the bionic hand consists of 27 different 3D-printed parts. The second set of assembly parts were purchased, these are components like ball bearings, gears, motors, the muscle sensors and the controller. In sum, 22 different components were purchased for this project. The last assembly parts are 159 screw-elements, like nuts, shells and washers.

B. Gripping Styles

Typical daily grips have been programmed and adjusted to be controlled with muscle sensors. The bionic hands 15 predefined grips use force feedback to ensure a tight grip on the taken object. Figure 13 displays a map of the predefined grips. The hand positions indicate the available functions in the function 1, function 2 or function 3 layers. The advantage of independent fingers makes it possible to create all kind of grips. Therefore, some functions close the index-finger for locating the object in the hand before the other fingers close. Another special grip is the anti-slip grip, using the small finger to prevent slippery objects from sliding out of the hand. Other objects require the parallel grips which close all finger in a parallel formation. To pick up thin objects like a tissue pack, the precision grips have been implemented. They move only the thumb and index-finger, the wearer decides which will be the moving finger. The keyboard grip enables the index-finger to point at something or to press a button. It is also possible to close this finger to activate the key grip, which is perfect for taking a ticket from a parking ticket machine. Small objects do not need contact to every finger, which makes moving them unnecessary. Therefore the tripod grips moves only the thumb, index- and middle finger. The ”pen” grip is a particularly advancement, enabling the wear to use a pen for drawing or writing. The bionic hand is a device that can be used to get back to your hobbies. Therefore, the extreme grips have been designed. They use

the edge of the finger tips to pick up resistors, cables or nails. The hook grip is a more robust grip for lifting heavy objects of up to three and a half kilograms, because the fingers are aligned to counterbalance the weight.

C. Scenarios

The 15 programmed gripping styles were tested in different scenarios, which show the application range of the bionic hand. Figure 15 shows different simple gripping examples, without an interaction of the left human hand, carried out by muscle contraction of the right forearm. Other examples with a left human hand interaction are shown in Figure 14. An important specification for grabbing an object is the closed loop control of the linear actuators, explained in Chapter IV-B.

1) Simple Grips without Human Interaction:

a) Normal: The first example shows the normal grip, used to grab the cap of a can. The big advantage in this example is the closed loop control, by which the finger movement of the bionic hand will automatically stop after grabbing the object.

b) Precision: Example (b) shows a precision grip, where the inside of the forefinger touches the front side of the thumb.

c) Normal: The third example shows the normal grip again. Grabbing a small ball is a good example for demonstrating the finger positions. Each finger will move as long as the actuator load is lower than the determined value. Therefore, it is possible to grab objects of different shapes such as a ball.

d) Precision: This scenario shows the same precision grip again, this time with a fragile object.

e) Pen: To fix a pen between the forefinger, middle finger and the thumb, the pen grip can be used. Because of the integrated 6-joint linkage, combined with the self-locking actuators, the mechanical construction of the fingers is stable enough to perform a safe mount of a pen.

f) Hook: The hook grip can be used to lift heavy objects like toolboxes or shopping bags, with a maximum weight of three and a half kilograms. The self-locking actuators will prevent an inadvertent finger movement, while grabbing. Example (f) shows a toolbox whit a total weight of 3.5kg.

2) Simple Grips with Human Hand Interaction:

a) Precision: This example shows a match, fixed between the forefinger and the thumb of the bionic hand. The difficult part in this example starts when the wearer attempts to light the match with the matchbox. At this point, some additional forces act on the match, and therefore it is possible that the match slips away. This example demonstrates that the mechanical requirements are given, to fix an object with two fingers safely without it slipping when an external force interacts.

b) Precision: In this example another precision grip, where the fingertips are touching is used. This grip can be used to fix small objects like a resistor or a paper.

c) Anti-Slip Normal: It is possible to enter the rotation mode from each grip. An integrated wrist rotation will replace the forearm rotation of a human hand. The anti-slip normal grip is a special grip for objects like a bottle. The little finger is in a retracted position and therefore it prevents objects from slipping through. After closing the fingers, the bottle is fixed enough to open the cap with the left human hand. To fill the liquid into a glass, the rotation mode can be activated to rotate the wrist for approximately 90 degrees.

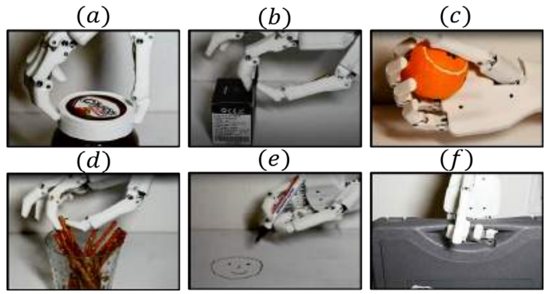


Fig. 13. Overview of 15 programmed gripping styles

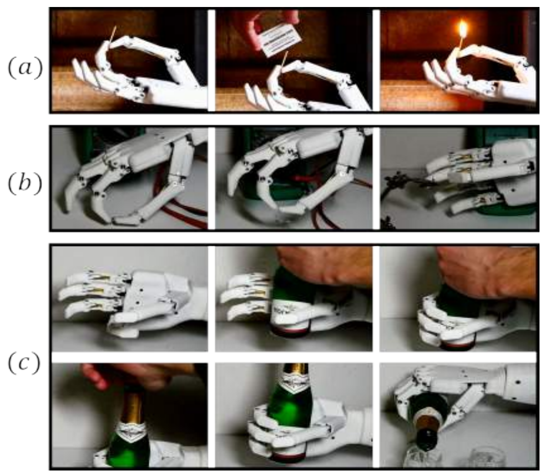


Fig. 14. Difficult gripping examples with an interaction of a left human hand

VI. CONCLUSIONS

The methods demonstrated have been used to develop a bionic hand using linear actuators to move five independent fingers. For increased usability, a rotating wrist and a motorized slewing thumb were implemented. Typical daily grips have been programmed, tested and adjusted to be controlled with on-skin muscle sensors. These sensors have been placed

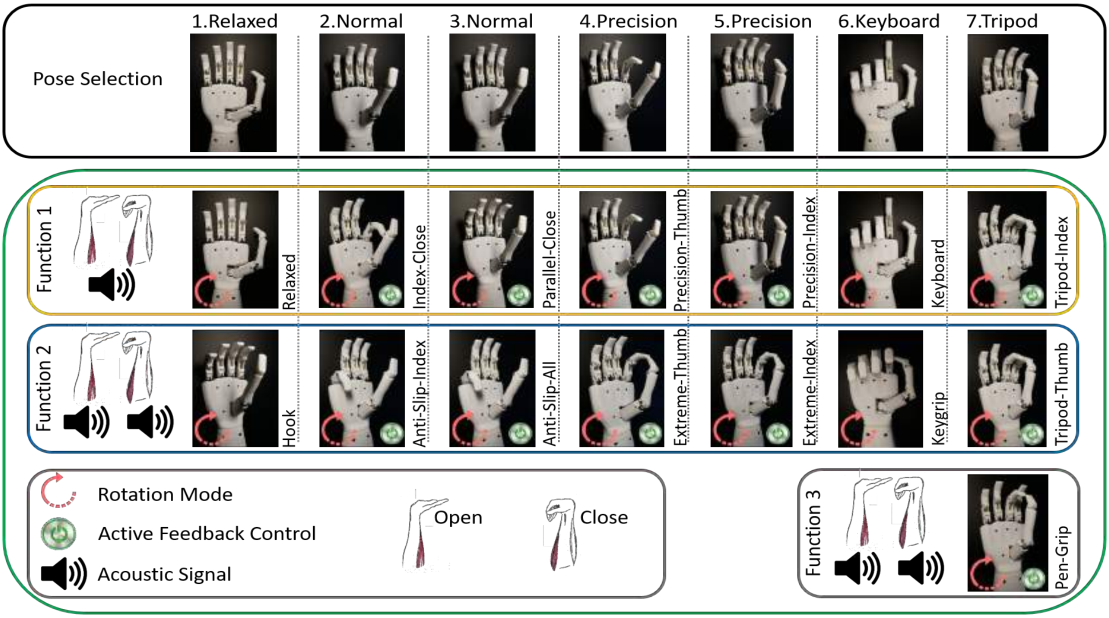


Fig. 15. Different daily lives gripping examples

on the forearm muscles, enabling the wearer to switch between different grips, rotate the wrist and control the fingers precisely. Tests of these muscle sensors revealed that some additional noise cancelling was necessary for interpretation of the muscle contraction. Therefore, a digital filter with a low pass characteristic was only used to smooth the rising muscle signal, a falling edge remained unmodified. This enhanced the controllability of the bionic hand extremely, and fine finger movements are now possible. To ensure a thin palm, the linear actuators have been combined with a six joint finger-linkage. This created a defined relation between the actuator position and the position of the fingertip. Those couplings are realised with metal tie rods with little clearance to achieve a good repeatability. The thumbs linear actuator is mounted on a slewing finger base, which can be positioned by a servo motor. It is then possible to grab a bottle and hold it securely enough for it to be manipulated. A big advantage of this bionic hand is the use of a force feedback, which turned out to be very precise. A human hand rotates the hand by using the forearm, however the bionic hand copied this function with the integration of a servo motor which acts as a wrist and permits a rotation of 135. The combination of the actuators and servo motors made it possible to define 15 different grips, which use force feedback and are precise enough to hold a small resistor or a thin piece of paper between the finger tips. The functionality of a rotating wrist is implemented in every function and allows a fast object manipulation. The controller, mounted into the forearm of the bionic hand was programmed, and about 85 percent of a 32kb memory have been used. The load on the actuators has never exceeded 20 percent to avoid damage to a grasped object. Nevertheless, if the force control notifies resistance, a short muscle impulse enables the fingers to close incrementally.

References

[1] G. C. Matrone, C. Cipriani, M. C. Carrozza, and G. Magenes, "Real-time myoelectric control of a multi-fingered hand prosthesis using principal components analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 40, Jun 2012.

[2] (2018) Touch Bionics: I-Limb Ultra. [Online]. Available: http://www.touchbionics.com/products/active-prostheses/i-limb-ultra

[3] (2018) Ottobock Grip Patterns. [Online]. Available: http://bebionic.com/the_hand/grip_patterns

[4] (2014) BBC: The incredible human hand and foot. [Online]. Available: http://www.bbc.com/news/science-environment-26224631

[5] J. T. Belter, J. L. Segil, A. M. Dollar, and R. F. Weir, "Mechanical design and performance specifications of anthropomorphic prosthetic hands: a review," *J Rehabil Res Dev*, vol. 50, no. 5, pp. 599–618, 2013.

[6] I.-D. Nicolae and P.-M. Nicolae, "Denoising highly distorted small currents in an environment with variable levels of noise," in *Electromagnetic Compatibility & Signal/Power Integrity (EMCSI), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 299–304.

[7] K. Or, M. Tomura, A. Schmitz, S. Funabashi, and S. Sugano, "Interpolation control posture design for in-hand manipulation," in *System Integration (SII), 2015 IEEE/SICE International Symposium on*. IEEE, 2015, pp. 187–192.

[8] D. G. Stănescu, M. E. Ardeleanu, and A. C. Stan, "Designing, simulation and testing of low current passive filters used in the didactic activity," in *Modern Power Systems (MPS), 2017 International Conference on*. IEEE, 2017, pp. 1–4.

[9] (2016) Actuonix: Miniature Linear Motion Series (Datasheet). [Online]. Available: https://s3.amazonaws.com/actuonix/Actuonix+PQ12+Datasheet.pdf
[10] K. Sakata and H. Fujimoto, "Perfect tracking control of servo motor based on precise model with pwm hold and current loop," in *Power Conversion Conference-Nagoya, 2007. PCC'07*. IEEE, 2007, pp. 1612–1617.
[11] (2017) Metal Gear Servo MG90S: Metal gear with one bearing (Datasheet). [Online]. Available: https://s3.amazonaws.com/actuonix/Actuonix+PQ12+Datasheet.pdf

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Extension of the Action Verb Corpus for Supervised Learning

Matthias Hirschmanner[1], Stephanie Gross[2], Brigitte Krenn[2], Friedrich Neubarth[2], Martin Trapp[2], Michael Zillich[1] and Markus Vincze[1]

***Abstract*— The Action Verb Corpus (AVC) is a multimodal dataset of simple actions for robot learning. The extension introduced here is especially geared to supervised learning of actions from human motion data. Recorded are RGB-D videos of the test scene, grayscale videos from the user's perspective, human hand trajectories, object poses and speech utterances. The three actions TAKE, PUT and PUSH are annotated with labels for the actions in different granularity.**

I. INTRODUCTION

Future social robots will have to acquire new tasks and behaviors on the go through interaction with users. They need to understand scenes, natural language instructions and user motions. In order to learn new actions via imitation or verbal instructions, empirical human data is needed. We introduced the Action Verb Corpus (AVC) as a multimodal dataset with simple object manipulation actions inspired by early parent-infant communication [1]. The extension presented in this paper is focused on supervised learning for action recognition from human motion data.

Existing datasets for action recognition that provide skeleton tracking often use the Microsoft Kinect camera such as the NTU RGB+D dataset [2] or the Montalbano dataset [3]. The Kinect tracks the whole-body skeleton but lacks individual finger tracking. For the dataset provided by Marin, Dominio and Zanuttigh [4], the Kinect as well as the Leap Motion sensor were used to capture the joint positions of fingers for American Sign Language gestures.

The extension of the AVC is geared towards robotic learning of interaction with objects. The joint positions of the fingers and the object poses are tracked. The recorded manipulations of objects located on a table are annotated in two degrees of granularity. Coarse labels reflect how the users refer to the action (e.g., TAKE, PUT, PUSH). Fine labels split an action into more granular motion primitives (e.g., REACH, GRAB, MOVE OBJECT).

II. DATASET

The AVC is a multimodal dataset of simple actions for robot learning from demonstration. It was recorded from inexperienced users performing the simple actions TAKE, PUT and PUSH with different objects according to visual instructions. They were verbalizing what they were doing in German. For example, the user moves the bottle to the left side of the box and says, "Ich nehme die Flasche und stelle sie neben die Schachtel" ("I take the bottle and put it next to the box").

For the extension of the Action Verb Corpus, users experienced with the system performed the same three basic actions arbitrarily. These actions were annotated afterwards to be used for supervised learning for action recognition. This approach was chosen to obtain recordings with good tracking performance for training a machine learning model. We will use the dataset for action classification of simple actions from human motion data in order to provide the basis for robotic learning from demonstration.

A. Setup

In the basic setup, a box, a bottle and a can are positioned on a table. The user wears the Oculus Rift DK2 virtual reality headset with the Leap Motion sensor mounted on top of it. A Microsoft Kinect camera is directed at the table for object tracking. During data collection, the user moves the object on the table and describes the actions he/she is performing. The speech utterances are recorded. The setup can be seen in Fig. 1.

The Leap Motion is a stereo infrared camera constructed particularly for hand tracking. The provided software fits a hand model to the pair of captured images to retrieve the joint positions. It returns the joint position of the human hand down to the singular finger segments with sub-millimeter precision [5].

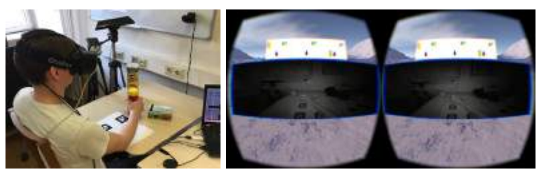


Fig. 1. The data collection setup with a user performing actions (left). Screenshot of the image shown in the Oculus Rift with the camera feed in the middle and the instructions on top (right).

The Oculus Rift headset provides the user's head pose. On the display of the headset, the user sees the scene in front of her/him as captured by the Leap Motion infrared cameras. This forces the user to direct the Leap Motion at the action she/he is performing. Therefore, the head pose can be used as an indication of gaze direction. It also ensures best possible hand tracking performance. The instructions the user has to

[1]Matthias Hirschmanner, Michael Zillich and Markus Vincze are with the Faculty of Electrical Engineering, Automation and Control Institute, Vision for Robotics, TU Wien, 1040 Wien, Austria {hirschmanner, zillich, vincze}@acin.tuwien.ac.at

[2] Stephanie Gross, Brigitte Krenn, Friedrich Neubarth and Martin Trapp are with the Austrian Research Institute for Artificial Intelligence (OFAI), 1010 Wien, Austria {stephanie.gross, brigitte.krenn, friedrich.neubarth, martin.trapp}@ofai.at

perform are displayed in the virtual reality headset above the camera images (Fig. 1).

Object tracking is performed on the monoscopic RGB images of the Kinect camera using an object tracker provided by the V4R library[1]. Models of the objects for tracking are created beforehand as described in [6]. Additionally, two binary features are saved: object is in contact with the table and object is in contact with a hand. The former is set automatically depending on the object's position, the latter is annotated manually. If the object is not in contact with a hand, averaging over consecutive object poses is performed weighted with the confidence of the object tracker because we assume the object does not move. This way, the jittering of the raw object-tracker data is reduced and occlusions do not impair tracking performance if the object was successfully tracked before.

The poses of the tracked entities (head, hands and objects) are transformed to a common coordinate frame and manually time-aligned.

B. Collected Data

The original Action Verb Corpus consists of 140 instances of TAKE/PUT actions and 110 instances of PUSH actions performed by 12 users following visual instructions. The focus is on word-object and word-action mapping.

The extension of the Action Verb Corpus consists of 210 instances of TAKE/PUT actions and 100 instances of PUSH actions performed by 2 experienced users without any instructions. The focus is on generating motion tracking data. A visualization of the tracked human arm and object poses is shown in Fig. 2. An issue in the original AVC is that the tracking information of the user's arm is lost sometimes while interacting with objects. An experienced user is able to operate the system in a way to get better tracking results and therefore more consistent data for a learning algorithm. The extension of the AVC is complementary to the original AVC.

The tracked data is annotated with action labels. Two types of annotations are created. The coarse annotation is how the user refers to the action. The classes of the coarse annotation are TAKE, PUT and PUSH. The fine annotation splits the actions into more granular motion primitives – REACH, GRAB, MOVE OBJECT and PLACE. The idea is that these primitives are more useful for the generation of robot actions while the coarse annotations reflect more complex motion concepts. For example, the robot might imitate human movement for reaching for an object. For grasping, it might switch to a different motion planner because the movement has to be adapted to the exact object pose. The coarse annotations are important for our overall goal of learning concepts of actions and link them with uttered verbs in order to acquire multimodal representations. This approach of labels with different granularity is similar to Koppula, Gupta and Saxena [7] who divide high level activities in sub-activities.

[1]https://www.acin.tuwien.ac.at/vision-for-robotics/software-tools/v4r-library/

The recordings of the extension of the Action Verb Corpus are represented by:

- 3D joint positions of the human arms, hands and fingers
- Head pose of the user
- Object poses with its corresponding confidence
- Binary features if the object touches a hand or the table
- Action annotations (coarse and fine)
- An animation of the tracked hands and objects
- RGB-D video of the scene
- Grayscale video from the user's perspective
- Recorded speech utterances

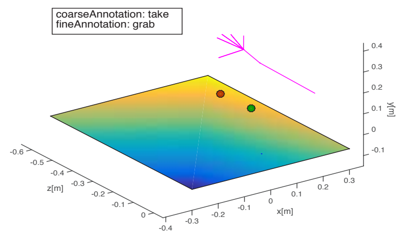


Fig. 2. Animation of the tracked data. A simplified version of the tracked arm is shown in magenta, the two objects are represented by the colored circles, the plane represents the table and the current action annotation is shown on top.

III. CONCLUSION AND FUTURE WORK

The Action Verb Corpus with its extension will be made available to the scientific community alongside this publication[2]. At the point of writing the dataset consists of 210 annotated TAKE/PUT and 100 PUSH actions. The data collection is still ongoing and will be further extended. We will use the dataset for action recognition of simple actions in order to provide the basis for robotic learning from demonstration. We want to extend the corpus with more complex actions. Additionally, we are working on alternative possibilities for human motion tracking that are less intrusive than our current setup. In a future step, a system will be implemented on a humanoid robot that will be able to detect different classes of actions and associate them with the user's utterance. Eventually, the robot should generate these actions and verbalize the imitated movements.

ACKNOWLEDGMENT

This research is supported by the Vienna Science and Technology Fund (WWTF), project RALLI – Robotic Action- Language Learning through Interaction (ICT15-045) and the CHIST-ERA project ATLANTIS (2287-N35). The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry for Science, Research and Economy.

[2]http://ralli.ofai.at/datasets.html

REFERENCES

[1] S. Gross, M. Hirschmanner, B. Krenn, F. Neubarth, and M. Zillich, "Action verb corpus," in *Proc. 2018 Language Recognition and Evaluation Conference*, Miyazaki, Japan, May 2018.

[2] A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang, "Ntu rgb+d: A large scale dataset for 3d human activity analysis," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[3] S. Escalera, X. Baró, J. Gonzalez, M. A. Bautista, M. Madadi, M. Reyes, V. Ponce-López, H. J. Escalante, J. Shotton, and I. Guyon, "Chalearn looking at people challenge 2014: Dataset and results," in *Workshop at the European Conference on Computer Vision*. Springer, 2014, pp. 459–473.

[4] G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with Leap Motion and Kinect devices," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1565–1569.

[5] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the Leap Motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.

[6] J. Prankl, A. Aldoma, A. Svejda, and M. Vincze, "RGB-D object modelling for object recognition and tracking," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 96–103.

[7] H. S. Koppula, R. Gupta, and A. Saxena, "Learning human activities and object affordances from RGB-D videos," *The International Journal of Robotics Research*, vol. 32, no. 8, pp. 951–970, 2013.

Philipp Zech, Justus Piater (Eds.)
Proceedings of the Austrian Robotics Workshop 2018
© 2018 *innsbruck* university press, ISBN 978-3-903187-22-1, DOI 10.15203/3187-22-1

Demand-driven Implementation of Human-Robot-Interaction in Manufacturing with Service Modelling

Kathleen Delang[1]*, Marcel Todtermuschke[1], Mohamad Bdiwi[1] and Matthias Putz[1]

***Abstract*— By combining advantages of humans and robots in the manufacturing process Human-Robot-Interaction (HRI) can solve many problems of todays production industry. Nevertheless, it still lacks industrial applications of this promising solution. The reasons are various and can be seen in uncertainties according to safety and a natural lower technical maturity of new systems. Another reason is the absence of a quantitative analysis of the benefits HRI can provide for the users. An assessment of existing work places as well as a selection and evaluation of potential improvements HRI may provide helps to justify investments. Therefore, a decision-making tool for investments in HRI will enlarge the number of use cases. This paper presents an approach to help producing companies comparing possibilities of HRI by evaluating existing process data.**

I. INTRODUCTION

The benefits of Human-Robot-Interaction (HRI) can be evaluated in an economic, ecological and social dimension covering acceptance and ergonomics. All mentioned dimensions combine different aspects. E.g. economy may be influenced by a higher flexibility, more added value, shorter tact time and the needed invest for the HRI system [1, p. 27]. These evaluation criteria of implemented HRI systems help to assess the potential of existing work places in advance. It is necessary to describe and demonstrate validated benefits of HRI according to individual motivation of a company [2].

Another reason for the lack of industrial applications of HRI is uncertainty in the context of safety regulations. Therefore ISO/TS 15066 [3] has been introduced in 2016. It defines allowed collision forces for different body parts. These specifications will help to build confidence for HRI systems in the whole process chain from technology providers, system integrators and end users [4] According to a shared workspace and the interaction during a performed task different forms of HRI can be distinguished reaching from coexistence to collaboration [5]. Thereby, the requirements for safety technology and the risk depend on the chosen form of interaction. Consequently, the necessary amount of money varies and the return of invest being the most important factor for investments in many companies [6, p. 518] depends on the level of interaction. For a methodology, assessing potentials of HRI a main requirement is flexibility since the developments in HRI are fast and latest trends have to be considered. The multi-layer approach of service modelling defines a meta-model with time-related process steps and an additional logical structure for the conditions and relations between predefined classes. Thereby, several models can be developed to achieve the overall objective with different methods. The presented approach provides the following benefits:

- Consideration of individual motivation
- Neutral selection and evaluation of work places
- Objective choice for the end user.

II. MULTI-LAYER APPROACH TO SERVICE MODELING

Modelling is a common solution in software development and helps structuring complex problems by defining architecture for a solution [7, p. 581]. The multi-layer approach of service modelling offers a flexible solution by designing a meta-model that defines requirements for different models to be applied in various applications [8], [9]. The context of the different layers is illustrated in Fig. 1.

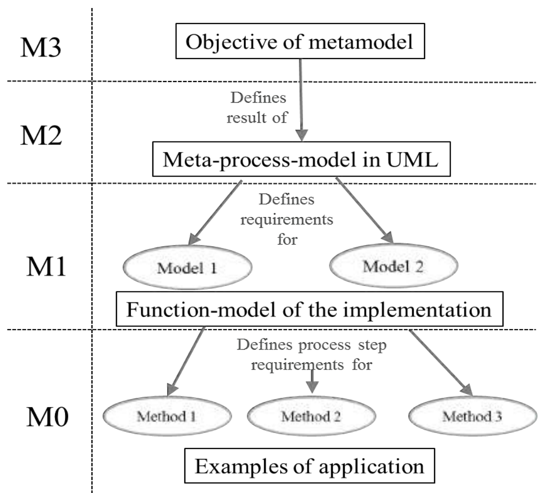


Fig. 1. Multi-layer approach with correlations (based on [8], [9])

The meta-model consists of a temporary structure for different process steps and a connected logical structure. In UML classes relate to possible attributes and are linked via associations and compositions [10, p.15]. The meta-model of the evaluation and selection of possible work places for HRI is presented in Fig. 2. Thereby, the initial data are the request

[1]All Authors are with the Fraunhofer Institute for Machine Tools and Forming Technology IWU, Chemnitz, Germany.
*Corresponding author: Kathleen Delang, phone: +49 371 5397-1265; kathleen.delang@iwu.fraunhofer.de

to change of the company and the given process parameters (marked bolted). The challenge is to structure the request and evaluate possible HRI benefits with the given process parameters.

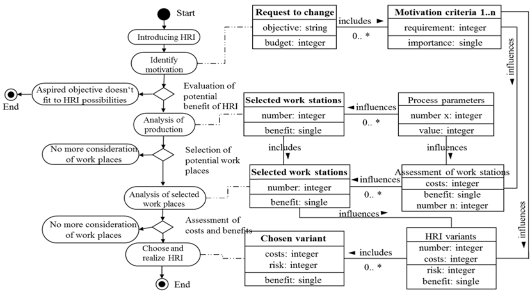


Fig. 2. Meta-model defining requirements for the planned models

Possible methods to address the needed objectives of each process step are shown in Fig. 3. In form of a morphological box the methods can be selected according to the needs and requirements of each user. Companies vary according to their available process data, the possibility to share data with external experts and their needed level of detailed analysis. The morphological box offers a set of possible tools to be chosen according to the individual constraints.

Thereby the structuring of the motivation is carried out first and a plant screening of suitable work places may follow to reduce the expense. The analysis of the production will be carried out according to the individual motivation and the selected work places with existing process data. For selected work places a cost-benefit-matrix or other evaluation methods are applied to compare possible HRI work places. As HRI can be designed in different levels this results in different concepts. Therefore, one work place may appear in the assessment with different HRI concepts.

The analysis of selected work places refers to individual motivation and requirements of the HRI end user. The best-rated work place is recommended to be realized since it provides the suitable benefit for the company. The realization is accomplished by risk sharing between technology providers with expertise in safety assessment, simulation or HRI concepts and integrators.

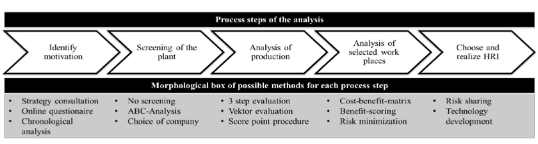


Fig. 3. Morphological box of suitable methods for each process step

III. FUTURE RESEARCH

To benefit from the presented approach a profile with advantages and requirements should be offered for each method of the morphological box. Thereby, companies can choose their individual most suitable methods and benefit from others experience. The profiles should provide an overview of necessary process data, provided benefits and level of detail for each method to simplify the choice. For an academic validation of the whole presented methodology a case study with at least three producing companies will be implemented.

ACKNOWLEDGMENT

Research supported by Federal Ministry for Economic Affairs and Energy, ZIM Zentrales Innovationsprogramm Mittelstand.

REFERENCES

[1] B. Mikus, U. Götze, and M. Schildt, "Kooperation zwischen mensch und roboter - ein beitrag zur nachhaltigen produktion?" *Der Betriebswirt*, vol. 2016, no. 2, pp. 25–31, 2016.

[2] K. Delang, M. Bdiwi, A.-K. Harsch, and M. Putz, "An approach of service modeling for the demand-driven implementation of human-robot-interaction in manufacturing," in *IEEE/RSJ International Conference onIntelligent Robots and Systems IROS2017, Workshop:Human-Robot Interaction in Collaborative Manufacturing Environments*, 2017.

[3] "ISO/TS 15066:2016: Robots and robotic devices collaborative robots," International Organization for Standardization, Geneva, CH, Standard, 2016.

[4] "Sicherheit in der Mensch-Roboter Kollaboration," Fraunhofer Austria & TÜV AUSTRIA Gruppe, Tech. Rep., 2016.

[5] M. Bdiwi, S. Krusche, and M. Putz, "Zone-based robot control for safe and efficient interaction between human and industrial robots," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA: ACM, 2017, pp. 83–84.

[6] M. Naumann, T. Dietz, and A. Kuss, "Mensch-maschine-interaktion," in *Industrie 4.0 in Produktion, Automatisierung und Logistik*, T. Bauernhansl, M. ten Hompel, and B. Vogel-Heuser, Eds. Springer Vieweg, 2014, ch. 25, pp. 508–529.

[7] C. Dickerson and D. Mavris, "A brief history of models and model based systems engineering and the case for relational orientation," vol. 7, pp. 581–592, 12 2013.

[8] H. Kern, M. Böttcher, S. Kühne, and K. Meyer, *Ansatz zur ganzheitlichen Erstellung und Verarbeitung von Dienstleistungsmodellen*. Heidelberg: Physica-Verlag HD, 2009, pp. 3–15.

[9] S. Grandt, "Entwicklung eines Referenzvorgehensmodells zur multikriteriellen Bewertung innovativer Sicherheitstechniken," PhD dissertation, Bergische Universität Wuppertal, 2015.

[10] B. Rumpe, *Modellierung mit UML: Sprache, Konzepte und Methodik*, ser. Xpert.press. Springer Berlin Heidelberg, 2011.

The sixth Austrian Robotics Workshop sought to bring together researchers, professionals and practitioners working on dierent topics in robotics to discuss recent developments and future challenges in robotics and its applications. The 2018 edition of the workshop series was be held at the University of Innsbruck in May 2018.

ISBN 978-3-903187-22-1

9 783903 187221